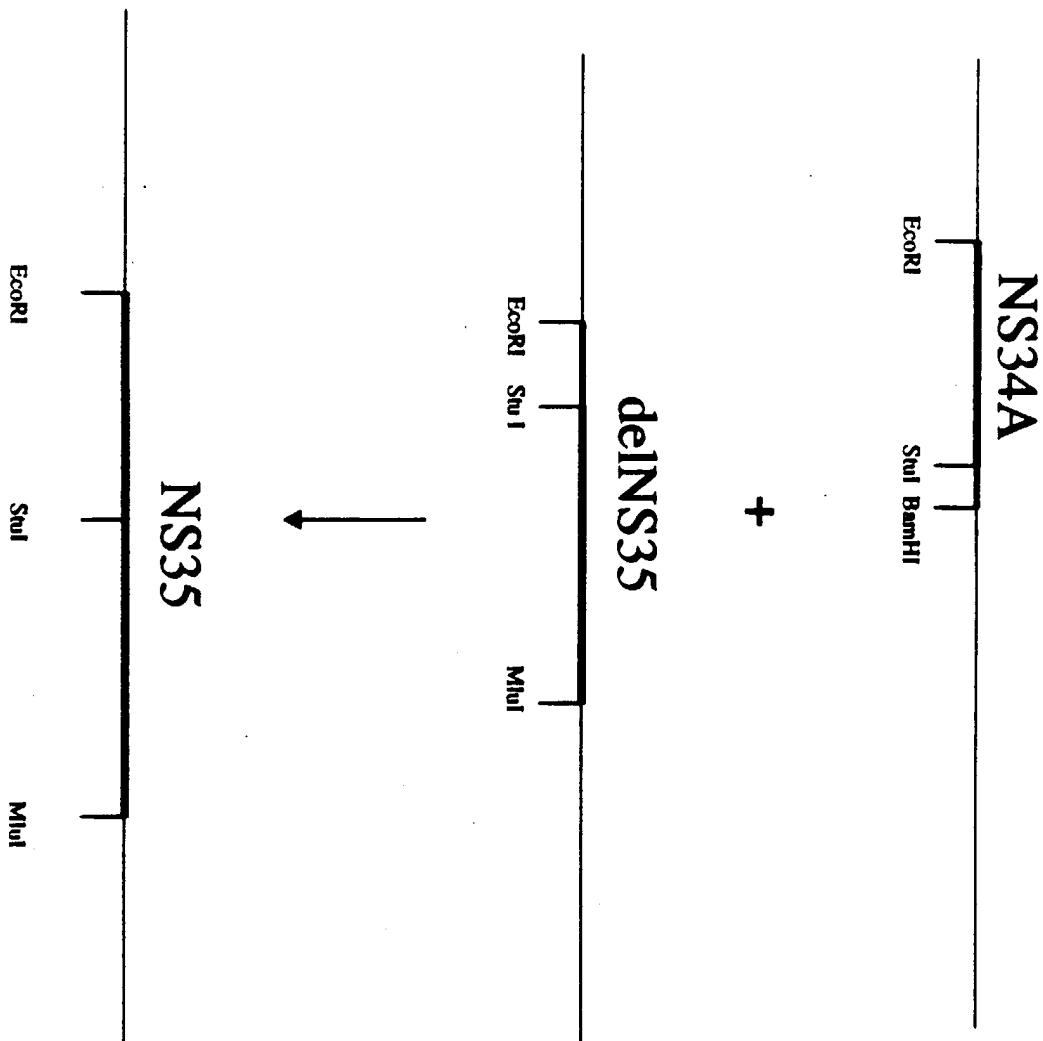
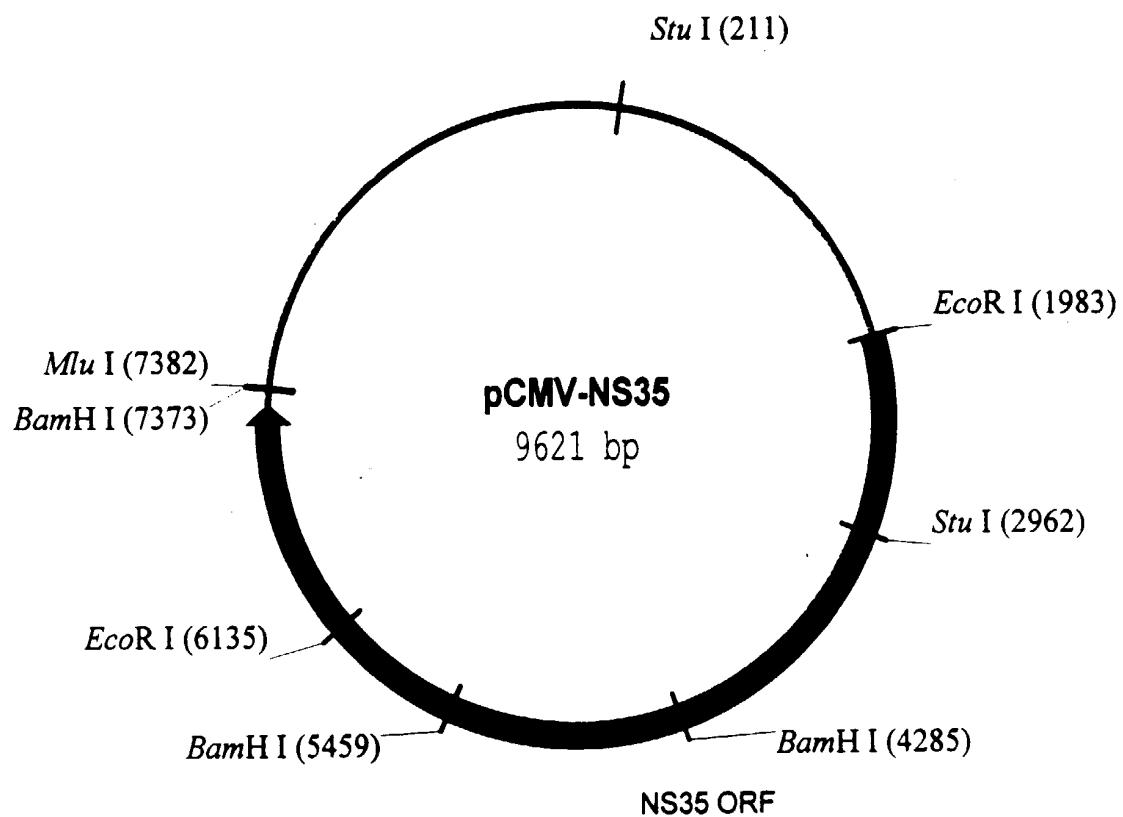


# Cloning Scheme for Generating pCMV-NS35

FIGURE 1



**FIGURE 2**



## FIGURE 3 - Page 1

1 14 TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG GAGACGGTCA CAGCTTGCT GTAGCGGAT  
AGCGCGCAAA GCCACTACTG CCACCTTTGG AGACTGTGTA CGTCGAGGGC CTCTGCCAGT GTCGAACAGA CATTGCGCTA

15 81 GCGGGGAGCA GACAAGCCCG TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTAACTATG CGGCATCAGA  
CGGCCCTCGT CTGTTCGGGC AGTCCCAGCG AGTCGCCCCAC AACCGCCCCAC AGCCCCGACC GAATTGATAC GCGTAGTCT

StuI

~~~~~

16 161 GCAGATTGTA CTGAGAGTGC ACCATATGAA GCTTTTGCA AAAGCCTAGG CCTCCAAAAAA AGCCTCCTCA CTACTTCTGG  
CGTCTAACAT GACTCTCACG TGGTATACTT CGAAAAACGT TTTCGGATCC GGAGGTTTT TCGGAGGAGT GATGAAGACC

17 241 AATAGCTCAG AGGCCGAGGC GGCTCGGGC TCTGCATAAA TAAAAAAAAT TAGTCAGCCA TGGGGCGGAG AATGGGGCGGA  
TTATCGAGTC TCCGGCTCCG CGGGAGCCGG AGACGTATTT ATTTTTTTA ATCAGTCGGT ACCCCGCCTC TTACCCGCCT

18 321 ACTGGGCGGG GAGGGAATTA TTGGCTATTG GCCATTGCAT ACGTTGTATC TATATCATAA TATGTACATT TATATTGGCT  
TGACCCGCC CTCCCTTAAT AACCGATAAC CGGTAACGTA TGCAACATAG ATATAGTATT ATACATGTAA ATATAACCGA

19 401 CATGTCCAAT ATGACCGCCA TGTTGACATT GATTATTGAC TAGTTATTAA TAGTAATCAA TTACGGGTC ATTAGTCAT  
GTACAGGTTA TACTGGCGGT ACAACTGTAA CTAATACTG ATCAATAATT ATCATTAGTT AATGCCAG TAATCAAGTA

20 481 AGCCCATATA TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC CCCGCCATT  
TCGGGTATAT ACCTCAAGGC GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC GGGTTGCTGG GGGCGGGTAA

21 561 GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTACGGT  
CTGCAGTTAT TACTGCATAC AAGGGTATCA TTGCGGTTAT CCCTGAAAGG TAACTGCAGT TACCCACCTC ATAAATGCCA

22 641 AAACTGCCCA CTGGCAGTA CATCAAGTGT ATCATATGCC AAGTCCGCC CCTATTGACG TCAATGACGG TAAATGGGCC  
TTTGACGGGT GAACCGTCAT GTAGTTCACA TAGTATACGG TTCAGGGGG GGATAACTGC AGTTACTGCC ATTTACCGGG

23 721 GCCTGGCATT ATGCCAGTA CATGACCTA CGGGACTTTC CTACTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC  
CGGACCGTAA TACGGGTAT GTACTGGAAT GCCCTGAAAG GATGAACCGT CATGTAGATG CATAATCAGT AGCGATAATG

24 801 CATGGTATG CGGTTTGCG AGTACACCAA TGGGCGTGA TAGCGTTTG ACTCACGGG ATTTCCAAGT CTCCACCCCA  
GTACCACTAC GCCAAACCG TCATGTGGTT ACCCGCACCT ATCGCAAAC TGAGTGGCCC TAAAGGTCA GAGGTGGGGT

25 881 TTGACGTCAA TGGGAGTTG TTTGGCACC AAAATCAACG GGACTTCCA AAATGTCGTAA ATAACCCCGC CCCGTTGACG  
AACTGCAGTT ACCCTCAAAC AAAACCGTGG TTTAGTTGC CCTGAAAGGT TTTACAGCAT TATTGGGCG GGGCAACTGC

26 961 CAAATGGCG GTAGGGTGT ACGGTGGAG GTCTATATAA GCAGAGCTCG TTAGTGAAC CGTCAGATCG CCTGGAGACG  
GTTTACCCGC CATCCGCACA TGCCACCTC CAGATATATT CGTCTCGAGC AAATCACTG GCAGTCTAGC GGACCTCTGC

27 1041 CCATCCACGC TGTGTTGACC TCCATAGAAAG ACACCGGGAC CGATCCAGCC TCCGCGGCCG GGAACGGTGC ATTGGAACGC  
GGTAGGTGCG ACAAAAGTGG AGGTATCTTC TGTGGCCCTG GCTAGGTGCG AGGCGCCGGC CCTTGCCACG TAACCTTGCG

28 1121 GGATTCCCCG TGCCAAGAGT GACGTAAGTA CCGCTATAG ACTCTATAGG CACACCCCTT TGGCTTTAT GCATGCTATA  
CCTAAGGGGC ACGGTTCTCA CTGCATTATC GGCGGATATC TGAGATATCC GTGTGGGAA ACCGAGAATA CGTACGATAT

29 1201 CTGTTTTGG CTTGGGGCCT ATACACCCCC GCTCCTTATG CTATAGGTGA TGGTATAGCT TAGCTATAG GTGTGGGTTA  
GACAAAAACC GAACCCCGGA TATGTGGGG CGAGGAATAC GATATCCACT ACCATATCGA ATCGGATATC CACACCCAT

30 1281 TTGACCATTA TTGACCACTC CCCTATTGGT GACGATACTT TCCATTACTA ATCCATAACA TGGCTTTG CCACAACTAT  
AACTGGTAAT AACTGGTGTAG GGGATAACCA CTGCTATGAA AGGTAATGAT TAGGTATTGT ACCGAGAAC GGTGTTGATA

31 1361 CTCTATTGGC TATATGCCAA TACTCTGTCC TTCAGAGACT GACACGGACT CTGTATTTT ACAGGATGGG GTCCATTAT  
GAGATAACCG ATATACGGTT ATGAGACAGG AAGTCTCTGA CTGTGCCCTGA GACATAAAA TGTCTACCC CAGGTAAATA

DRAFT

## FIGURE 3 - Page 2

1441 TATTTACAAA TTCACATATA CAAACAAGCC GTCCCCCGTG CCCGCAGTTT TTATTAACAA TAGCGTGGGA TCTCCGACAT  
ATAAAATGTTT AAGTGTATAT GTTGTGCGG CAGGGGGCAC GGGCGTCAA AATAATTGT ATCGCACCC AGAGGCTGTA

---

1521 CTCGGGTACG TGTTCGGAC ATGGGCTT CTCCGGTAGC GGCGGAGCTT CCACATCCGA GCCCTGGTCC CATCCGTCCA  
GAGCCCATGC ACAAGGCCTG TACCCGAGAA GAGGCCATCG CCGCCTCGAA GGTGTAGGCT CGGGACCAGG GTAGGCAGGT

---

1601 GCGGCTCATG GTCGCTCGGC AGCTCCTGTC TCCTAACAGT GGAGGCCAGA CTTAGGCACA GCACAATGCC CACCACCAC  
CGCCGAGTAC CAGCGAGCCG TCGAGGAACG AGGATTGTCA CCTCCGGTCT GAATCCGTGT CGTGTACGG GTGGTGGTGG

---

1681 AGTGTGCCGC ACAAGGCCGT GGCGGTAGGG TATGTGTCTG AAAATGAGCT CGGAGATTGG GCTCGCACCT GGACGCAGAT  
TCACACGGCG TGTTCCGGCA CGGCATCCC ATACACAGAC TTTTACTCGA GCCTCTAACCG CGAGCGTGGGA CCTGCCTCTA

---

1761 GGAAGACTTA AGGCAGCGGC AGAAGAAGAT GCAGGCAGCT GAGTTGTGT ATTCTGATAA GAGTCAGAGG TAACTCCCGT  
CCTCTGAAT TCCGTCGCCG TCTTCTCTA CGTCCGTGCA CTCAACAACA TAAGACTATT CTCAGTCTCC ATTGAGGGCA

---

1841 TGCGGTGCTG TTAACGGTGG AGGGCAGTGT AGTCTGAGCA GTACTCGTTG CTGCCCGCG CGCCACCAGA CATAATAGCT  
ACGCCACGAC AATTGCCACC TCCCCTCACA TCAGACTCGT CATGAGCAAC GACGGCGCGC GCGGTGGTCT GTATTATCGA

---

+2 M A A

EcoRI

1921 GACAGACTAA CAGACTGTTC CTTTCCATGG GTCTTTCTG CAGTCACCGT CGTCACCTA AGAATTCAAC ATGGCTGCAT  
CTGCTGATT GTCTGACAAG GAAAGTACG CAGAAAGAC GTCACTGGAT TCTTAAGTGG TACCGACGTA

---

+2 Y A A Q G Y K V L V L N P S V A A T L G F G A Y M S K  
2001 ATGCAGCTCA GGGCTATAAG GTGCTAGTAC TCAACCCCTC TGTTGCTGCA ACACTGGCT TTGGTGTCTTA CATGTCCAAG  
TACGTCGAGT CCCGATATTC CACGATCATG AGTGGGGAG ACAACGACGT TGTGACCCGA AACCACGAAT GTACAGGTT

---

+2 A H G I D P N I R T G V R T I T T G S P I T Y S T Y G  
2081 GCTCATGGGA TCGATCCTAA CATCAGGACC GGGGTGAGAA CAATTACAC TGGCAGCCCC ATCACGTACT CCACCTACGG  
CGAGTACCCCT AGCTAGGATT GTAGTCCCTGG CCCCACTCTT GTTAATGGTG ACCGTCGGGG TAGTGCATGA GGTGGATGCC

---

+2 K F L A D G G C S G G A Y D I I I C D E C H S T D A  
2161 CAAGTTCCCTT GCCGACGGCG GGTGCTCGGG GGGCGCTTAT GACATAATAA TTTGTGACGA GTGCCACTCC ACGGATGCCA  
GTTCAAGGAA CGGCTGCCGC CCACGAGCCC CCCCGGAATA CTGTATTATT AAACACTGCT CACGGTGAGG TGCCCTACGGT

---

+2 T S I L G I G T V L D Q A E T A G A R L V V L A T A T  
2241 CATCCATCTT GGGCATTGGC ACTGTCCTTG ACCAAGCAGA GACTGCGGGG GCGAGACTGG TTGTGCTCGC CACCGCCACC  
GTAGGTAGAA CCCGTAACCG TGACAGGAAC TGGTTCGTCT CTGACGCCCC CGCTCTGACC AACACGAGCG GTGGCGGTGG

---

+2 P P G S V T V P H P N I E E V A L S T T G E I P F Y G  
2321 CCTCCGGGCT CCGTCACTGT GCCCCATCCC AACATCGAGG AGGTTGCTCT GTCCACCCACC GGAGAGATCC CTTTTACGG  
GGAGGCCCCGA GGCAGTGACA CGGGTAGGG TTGTAGCTCC TCCAACGAGA CAGGTGGTGG CCTCTCTAGG GAAAATGCC

---

+2 K A I P L E V I K G G R H L I F C H S K K K C D E L  
2401 CAAGGCTATC CCCCTCGAAG TAATCAAGGG GGGGAGACAT CTACATCTCT GTCAATTCAA GAAGAAGTGC GACGAACCTCG  
GTTCCGATAG GGGGAGCTTC ATTAGTCTCC CCCCTCTGTA GAGTAGAAGA CAGTAAGTT CTTCTTCAGC CTGCTTGAGC

---

+2 A A K L V A L G I N A V A Y Y R G L D V S V I P T S G  
2481 CCGCAAAGCT GGTCGCATG GGCATCAATG CCGTGGCCTA CTACCGCGGT CTTGACGTGT CCGTCATCCC GACCAGCGGC  
GGCGTTCGA CCAGCGTAAC CCGTAGTTAC GGCACCGGAT GATGGCGCCA GAACTGCACA GGCAGTAGGG CTGGTCCGGC

---

+2 D V V V V A T D A L M T G Y T G D F D S V I D C N T C  
2561 GATGTTGTCG TCGTGGCAAC CGATGCCCTC ATGACCGGGCT ATACCGGCA CTTGACACTCG GTGATAGACT GCAATACGTG  
CTACAACAGC AGCACCGTTG GCTACGGAG TACTGGCGA TATGGCCGCT GAAGCTGAGC CACTATCTGA CGTATGAC

## FIGURE 3 - Page 3

+2 V T Q T V D F S L D P T F T I E T I T L P Q D A V S  
 2641 . TGTCACCCAG ACAGTCGATT TCAGCCTGGA CCCTACCTTC ACCATTGAGA CAATCACGCT CCCCAAGAT GCTGTCCTCC  
 ACAGTGGTC TGTCAGCTAA AGTCGGAAGT GGATGGAAG TGGTAACCTCT GTAGTGCAG GGGGGTTCTA CGACAGAGGG

---

+2 R T Q R R G R T G R G K P G I Y R F V A P G E R P S G  
 2721 GCACTCAACG TCGGGGCAGG ACTGGCAGGG GGAAGCCAGG CATCTACAGA TTTGTGGCAC CGGGGGAGCG CCCCTCCGGC  
 CGTGAGTTGC AGCCCCGTCC TGACCGTCCC CCTTCGGTCC GTAGATGTCT AACACCGTG GCCCCCTCGC GGGGAGGCCG

---

+2 M F D S S V L C E C Y D A G C A W Y E L T P A E T T V  
 2801 ATGTTCAACT CGTCCGTCT CTGTGAGTGC TATGACGCAG GCTGTGCTTG GTATGAGCTC ACGCCCGCCG AGACTACAGT  
 TACAAGCTGA GCAGGCAGGA GACACTCACG ATACTGCGTC CGACACGAAC CATACTCGAG TGCGGGCGGC TCTGATGTCA

---

+2 R L R A Y M N T P G L P V C Q D H L E F W E G V F T  
 StuI  
 ~~  
 2881 TAGGCTACGA GCGTACATGA ACACCCCCGGG GCTTCCCGTG TGCCAGGACC ATCTGAATT TTGGGAGGGC GTCTTACAG  
 ATCCGATGCT CGCATGTACT TGTGGGGCCC CGAAGGGCAC ACGGTCTGG TAGAACTTAA AACCTCCCG CAGAAATGTC

---

+2 G L T H I D A H F L S Q T K Q S G E N L P Y L V A Y Q  
 StuI  
 ~~~  
 2961 GCCTCACTCA TATAGATGCC CACTTTCTAT CCCAGACAAA GCAGAGTGGG GAGAACCTTC CTTACCTGGT AGCGTACCAA  
 CGGAGTGAGT ATATCTACGG GTGAAAGATA GGGTCTGTT CGTCTCACCC CTCTGGAAG GAATGGACCA TCGCATGGTT

---

+2 A T V C A R A Q A P P P S W D Q M W K C L I R L K P T  
 3041 GCCACCGTGT GCGCTAGGGC TCAAGCCCT CCCCCATCGT GGGACCAGAT GTGAAAGTGT TTGATTCGCC TCAAGCCAC  
 CGGTGGCACA CGCGATCCCG AGTCGGGGA GGGGTAGCA CCCTGGTCTA CACCTCACCA AACTAAGCGG AGTTGGGTG

---

+2 L H G P T P L L Y R L G A V Q N E I T L T H P V T K  
 3121 CCTCCATGGG CCAACACCCC TGCTATACAG ACTGGCGCT GTTCAGAATG AAATCACCT GACGCACCA GTCACCAAAT  
 GGAGGTACCC GGTTGTGGGG ACGATATGTC TGACCCGCGA CAAGTCTTAC TTTAGTGGGA CTGCGTGGGT CAGTGGTTA

---

+2 Y I M T C M S A D L E V V T S T W V L V G G V L A A L  
 3201 ACATCATGAC ATGCATGTCG GCCGACCTGG AGGTGTCAC GAGCACCTGG GTGCTCGTT GCGGCCTCCT GGCTGCTTTG  
 TGTAGTACTG TACGTACAGC CGGCTGGACC TCCAGCAGTG CTCGTGGACC CACGAGCAAC CGCCGCAAGGA CGGACGAAAC

---

+2 A A Y C L S T G C V V I V G R V V L S G K P A I I P D  
 3281 GCCGCGTATT GCCTGTCAAC AGGCTGCGTG GTCATAGTGG GCAGGGTCGT CTTGTCCGGG AAGCCGGCAA TCATACCTGA  
 CGGCGCATAA CGGACAGTTG TCCGACGCAC CAGTATCACC CGTCCCAGCA GAACAGGCCCTC TCGGGCCGTT AGTATGGACT

---

+2 R E V L Y R E F D E M E E C S Q H L P Y I E Q G M M  
 3361 CAGGGAAGTC CTCTACCGAG AGTCGATGA GATGGAAGAG TGCTCTCAGC ACTTACCGTA CATCGAGCAA GGGATGATGC  
 GTCCCTCAG GAGATGGCTC TCAAGCTACT CTACCTCTC ACGAGAGTCG TGAATGGCAT GTAGCTCGTT CCCTACTACG

---

+2 L A E Q F K Q K A L G L L Q T A S R Q A E V I A P A V  
 3441 TCGCCGAGCA GTTCAAGCAG AAGGCCCTCG GCCTCCTGCA GACCGCGTCC CGTCAGGCAG AGGTTATCGC CCCTGCTGTC  
 AGCGGCTCGT CAAGTTCGTC TTCCGGGAGC CGGAGGACGT CTGGCGCAGG GCAGTCCGTC TCCAATAGCG GGGACGACAG

---

+2 Q T N W Q K L E T F W A K H M W N F I S G I Q Y L A G  
 3521 CAGACCAACT GGCAAAACT CGAGACCTTC TGGCGAAGC ATATGTGGAA CTTCATCAGT GGGATACAAT ACTTGGCGGG  
 GTCTGGTGA CCGTTTTGA GCTCTGGAA ACCCGCTTCG TATACACCTT GAAGTAGTCA CCCTATGTTA TGAACCGCCC

---

+2 L S T L P G N P A I A S L M A F T A A V T S P L T T  
 3601 CTTGTCAACG CTGCTGGTA ACCCCGCCAT TGCTTCATTG ATGGCTTTA CAGCTGCTGT CACCAAGCCCA CTAACCACTA  
 GAACAGTTGC GACGGACCAT TGGGGCGGT ACGAAGTAAC TACCGAAAAT GTCGACGACA GTGGTGGGT GATTGGTGT

---

+2 S Q T L L F N I L G G W V A A Q L A A P G A A T A F V  
 3681 GCCAAACCCCT CCTCTTCAAC ATATTGGGGG GGTGGGTGGC TGCCCGCTC GCCGCCCGGTG TGCCCGCTAC TGCCTTGTG  
 CGGTTGGGA GGAGAAGTTG TATAACCCCC CCACCCACCG ACGGGTCGAG CGGGGGGGC CACGGCGATG ACGGAAACAC

## FIGURE 3 - Page 4

+2 G A G L A G A A I G S V G L G K V L I D I L A G Y G A  
 3761 GGCCTGGCT TAGCTGGCGC CGCCATCGC AGTGTGGAC TGGGGAAAGGT CCTCATAGAC ATCCTTGAG GGTATGGCGC  
 CCGCGACCGA ATCGACCGCG GCGGTAGCCG TCACAACTG ACCCCTTCCA GGAGTATCTG TAGGAACGTC CCATACCGCG

+2 G V A G A L V A F K I M S G E V P S T E D L V N L L  
 3841 GGGCGTGGCG GGAGCTCTTG TGGCATCAA GATCATGAGC GGTGAGGTCC CCTCCACGGA GGACCTGGTC AATCTACTGC  
 CCCGCACCGC CCTCGAGAAC ACCGTAAGTT CTAGTACTCG CCACTCCAGG GGAGGTGCCT CCTGGACCAAG TTAGATGACG

+2 P A I L S P G A L V V G V V C A A I L R R H V G P G E  
 3921 CCGCCATCCT CTCGCCCCGGA GCCCTCGTAG TCGCGTGGT CTGTGCAGCA ATACTGCGCC GGACAGTTGG CCCGGCGAG  
 GCGGGTAGGA GAGCGGGCCT CGGGAGCATC AGCCGCACCA GACACGTCGT TATGACGCGG CCGTGCACCC GGGCCCGCTC

+2 G A V Q W M N R L I A F A S R G N H V S P T H Y V P E  
 4001 GGGGCAGTGC AGTGGATGAA CCGGCTGATA GCCTTCGCT CCCGGGGAA CCATGTTCC CCCACGCACT ACGTGCCGGA  
 CCCCGTCACG TCACCTACTT GGCGACTAT CGGAAGCGGA GGGCCCTT GGTACAAAGG GGGTGCCTGA TGCACGGCCT

+2 S D A A A R V T A I L S S L T V T Q L L R R L H Q W  
 4081 GAGCGATGCA GCTGCCCGCG TCACTGCCAT ACTCAGCAGC CTCACTGTAA CCCAGCTCCT GAGGCAGCTG CACCAAGTGG  
 CTCGCTACGT CGACGGCGC AGTGCAGGTA TGAGTCGTCG GAGTGACATT GGGTCGAGGA CTCCGCTGAC GTGGTCACCT

+2 I S S E C T T P C S G S W L R D I W D W I C E V L S D  
 4161 TAAGCTCGGA GTGTACCACT CCATGCTCCG GTTCTGGCT AAGGGACATC TGGGACTGGA TATGCAGGT GTTGAGCGAC  
 ATTCGAGCCT CACATGGTGA GGTACGAGGC CAAGGACCGA TTCCCTGTAG ACCCTGACCT ATACGCTCCA CAACTCGCTG

+2 F K T W L K A K L M P Q L P G I P F V S C Q R G Y K G  
 BamHI  
 -----

4241 TTTAAGACCT GGCTAAAAGC TAAGCTCATG CCACAGCTGC CTGGGATCCC CTTTGTGTCC TGCCAGCGCG GGTATAAGGG  
 AAATTCTGGA CCGATTTCG ATTCGAGTAC GGTGTCGACG GACCCTAGGG GAAACACAGG ACGGTCGCGC CCATATTCCC

+2 V W R G D G I M H T R C H C G A E I T G H V K N G T  
 4321 GGTCTGGCGA GGGGACGGCA TCATGCACAC TCGCTGCCAC TGTGGAGCTG AGATCACTGG ACATGTCAA ACGGGACGA  
 CCAGACCGCT CCCCTGCCGT AGTACGTGTG AGCGACGGTG ACACCTCGAC TCTAGTGACC TGTACAGTT TTGCCCTGCT

+2 M R I V G P R T C R N M W S G T F P I N A Y T T G P C  
 4401 TGAGGATCGT CGGTCTAGG ACCTGCAGGA ACATGTGGAG TGGGACCTTC CCCATTAATG CCTACACCAC GGGCCCTGT  
 ACTCCTAGCA GCCAGGATCC TGGACGTCC TGTACACCTC ACCCTGGAAG GGGTAATTAC GGATGTGGTG CCCGGGGACA

+2 T P L P A P N Y T F A L W R V S A E E Y V E I R Q V G  
 4481 ACCCCCCCTTC CTGCGCCGAA CTACACGTT GCGCTATGGA GGGTGTCTGC AGAGGAATAC GTGGAGATAA GGCAGGTGGG  
 TGGGGGAAG GACCGGGCTT GATGTGCAAG CGCGATACCT CCCACAGACG TCTCCTTATG CACCTCTATT CCGTCCACCC

+2 D F H Y V T G M T T D N L K C P C Q V P S P E F F T  
 4561 GGACTTCCAC TACGTGACGG GTATGACTAC TGACAATCTT AAATGCCGTG GCCAGGTCCC ATCGCCGAA TTTTCACAG  
 CCTGAAGGTG ATGCACTGCC CATACTGATG ACTGTTAGAA TTTACGGCA CGGTCAGGG TAGCGGGCTT AAAAGTGTG

+2 E L D G V R L H R F A P P C K P L L R E E V S F R V G  
 4641 AATTGGACGG GGTGCGCCTA CATAGGTTG CGCCCCCTG CAAGCCCTG CTGGGGAGG AGGTATCATT CAGAGTAGGA  
 TTAACCTGCC CCACCGGGAT GTATCCAAAC GCGGGGGGAC GTTCGGGAAC GACGCCCTCC TCCATAGTAA GTCTCATCCT

+2 L H E Y P V G S Q L P C E P E P D V A V L T S M L T D  
 4721 CTCCACGAAT ACCCGGTAGG GTCGCAATT CCTTGCGAGC CCGAACCGGA CGTGGCCGTG TTGACGTCCA TGCTCACTGA  
 GAGGTGCTTA TGGGCCATCC CAGCGTTAAT GGAACGCTCG GGCTTGGCT GCACCGGCAC AACTGCAGGT ACGAGTGACT

+2 P S H I T A E A A G R R L A R G S P P S V A S S S A  
 4801 TCCCTCCCAT ATAACAGCAG AGGCAGGCCGG GCGAACGGTG GCGAGGGAT CACCCCCCTC TGTGGCCAGC TCCTCGGCTA  
 AGGGAGGGTA TATTGTCGTC TCCGCGGCC CGCTTCCAAC CGCTCCCTA GTGGGGGAG ACACCGGTG AGGAGCCGAT

## FIGURE 3 - Page 5

+2 S Q L S A P S L K A T C T A N H D S P D A E L I E A N  
 4881 GCCAGCTATC CGCTCCATCT CTCAGGCAA CTTGCACCGC TAACCATGAC TCCCCGTGATG CTGAGCTCAT AGAGGCCAAC  
 CGGTCGATAG GCGAGGTAGA GAGTTCCGTT GAACGTGGCG ATTGGTACTG AGGGGACTAC GACTCGAGTA TCTCCGGTTG

---

+2 L L W R Q E M G G N I T R V E S E N K V V I L D S F D  
 4961 CTCCTATGGA GGCAGGAGAT GGGCGGCAAC ATCACCGAGG TTGAGTCAGA AAACAAAGTG GTGATTCTGG ACTCCTCGA  
 GAGGATACCT CCGTCCTCTA CCCGCCGTTG TAGTGGTCCC AACTCAGTCT TTTGTTTCAC CACTAAGACC TGAGGAAGCT

---

+2 P L V A E E D E R E I S V P A E I L R K S R R F A Q  
 5041 TCCGCTGTG GCGGAGGAGG ACGAGCGGA GATCTCCGTA CCCGCAGAAA TCCTCGGAA GTCTCGGAGA TTCGCCAGG  
 AGGCGAACAC CGCCTCTCC TGCTCGCCCT CTAGAGGCAT GGGCGTCTT AGGACGCCCT CAGAGCCTCT AAGCGGGTCC

---

+2 A L P V W A R P D Y N P P L V E T W K K P D Y E P P V  
 5121 CCCTGCCGT TTGGCGCGG CCGGACTATA ACCCCCCGCT AGTGGAGACG TGGAAAAAGC CCGACTACGA ACCACCTGTG  
 GGGACGGGCA AACCCGCGCC GGCCTGATAT TGGGGGGCGA TCACCTCTGC ACCTTTTCG GGCTGATGCT TGGTGGACAC

---

+2 V H G C P L P P P K S P P V P P P R K K R T V V L T E  
 5201 GTCCATGGCT GCCCCTTCC ACCTCCAAAG TCCCCTCCTG TGCCCTCCCGC TCGGAAGAAG CGGACGGTGG TCCTCACTGA  
 CAGGTACCGA CGGGCGAAGG TGGAGGTTTC AGGGGAGGAC ACGGAGGCGG AGCCTCTTC GCCTGCCACC AGGAGTGA

---

+2 S T L S T A L A E L A T R S F G S S S T S G I T G D  
 5281 ATCAACCTA TCTACTGCCT TGGCGAGCT CGCCACCAGA AGCTTGGCA GCTCTCAAC TTCCGGCATT ACGGCGACA  
 TAGTTGGAT AGATGACGGA ACCGGCTCGA GCGGTGGTCT TCGAAACCGT CGAGGAGTTG AAGGCCGTA TGCCCGCTGT

---

+2 N T T T S S E P A P S G C P P D S D A E S Y S S M P P  
 5361 ATACGACAAC ATCCTCTGAG CCCGCCCTT CTGGCTGCC CCCCAGTCC GACGCTGAGT CCTATTCTC CATGCCCCCC  
 TATGCTGTTG TAGGAGACTC GGGCGGGAA GACCGACGGG GGGGCTGAGG CTGCGACTCA GGATAAGGAG GTACGGGGGG

---

+2 L E G E P G D P D L S D G S W S T V S S E A N A E D V  
 BamHI  
 ~~~~~~

5441 CTGGAGGGGG AGCCTGGGA TCCGGATCTT AGCGACGGGT CATGGTCAAC GGTCAGTAGT GAGGCCAACG CGGAGGATGT  
 GACCTCCCCC TCGGACCCCT AGGCCTAGAA TCGCTGCCA GTACCAGTTG CCAGTCATCA CTCCGGTTGC GCCTCCTACA

---

+2 V C C S M S Y S W T G A L V T P C A A E E Q K L P I  
 5521 CGTGTGCTGC TCAATGTCTT ACTCTTGAC AGGCGCACTC GTCACCCCGT GCGCCCGGA AGAACAGAAA CTGCCATCA  
 GCACACGACG AGTTACAGAA TGAGAACCTG TCCCGTGAG CAGTGGGGCA CGCGCGCCT TCTGTCTTT GACGGTAGT

---

+2 N A L S N S L L R H H N L V Y S T T S R S A C Q R Q K  
 5601 ATGCACTAAG CAACTCGTTG CTACGTCAAC ACAATTGGT GTATTCCACC ACCTCACGCA GTGCTGCCA AAGGCAGAAC  
 TACGTGATTC GTTGAGCAAC GATGCAGTGG TGTAAACCA CATAAGGTGG TGGAGTGCCT CACGAACGGT TTCCGTCTTC

---

+2 K V T F D R L Q V L D S H Y Q D V L K E V K A A A S K  
 5681 AAAGTCACAT TTGACAGACT GCAAGTTCTG GACAGCCATT ACCAGGACGT ACTCAAGGAG GTTAAAGCAG CGCGCTCAAA  
 TTTCAGTGTAA AACTGTCTGA CGTTCAAGAC CTGTCGGTAA TGGCCTGCA TGAGTTCTC CAATTGTC GCGCAGTT

---

+2 V K A N L L S V E E A C S L T P P H S A K S K F G Y  
 5761 AGTGAAGGCT AACTTGCTAT CCGTAGAGGA AGCTTGACGC CTGACGCCCA CACACTCAGC CAAATCCAAG TTTGGTTATG  
 TCACTCCGA TTGAACGATA GGCATCTCT TCGAACGTG GACTGCGGGG GTGTGAGTCG GTTTAGGTT AAACCAATAC

---

+2 G A K D V R C H A R K A V T H I N S V W K D L L E D N  
 5841 GGGCAAAAGA CGTCCGTTGC CATGCCAGAA AGGCGTAAC CCACATCAAC TCCGTGTTGA AAGACCTCT GGAAGACAAT  
 CCCGTTTCT GCAGGCAACG GTACGGTCTT TCCGGCATTG GGTGTAGTTG AGGCACACCT TTCTGGAAAGA CCTTCTGTTA

---

+2 V T P I D T T I M A K N E V F C V Q P E K G G R K P A  
 5921 GTAACACCAA TAGACACTAC CATCATGGCT AAGAACGAGG TTTCTGCGT TCAGCCTGAG AAGGGGGTC GTAAGCCAGC  
 CATTGTGGTT ATCTGTGATG GTAGTACCGA TTCTGCTCC AAAAGACGCA AGTCGGACTC TTCCCCCAG CATTGGTCG

## FIGURE 3 - Page 6

+2 R L I V F P D L G V R V C E K M A L Y D V V T K L P  
 6001 TCGTCTCATC GTGTTCCCCG ATCTGGCGT GCGCGTGTG GAAAAGATGG CTTTGTACGA CGTGGTTACA AAGCTCCCT  
 AGCAGAGTAG CACAAGGGGC TAGACCCGCA CGCGCACACG CTTTCTACC GAAACATGCT GCACCAATGT TTCGAGGGGA

---

+2 L A V M G S S Y G F Q Y S P G Q R V E F L V Q A W K S  
 EcorI  
 ~~~~~~

6081 TGGCCGTGAT GGGAAAGCTCC TACGGATTCC AATACTCACC AGGACAGCGG GTTGAATTCC TCGTGCAGC GTGGAAGTCC  
 ACCGGCACTA CCCTTCGAGG ATGCCTAAGG TTATGAGTGG TCCGTGCGC CAACTTAAGG AGCACGTTCG CACCTTCAGG

---

+2 K K T P M G F S Y D T R C F D S T V T E S D I R T E E  
 6161 AAGAAAACCC CAATGGGGTT CTCGTATGAT ACCCGCTGCT TTGACTCCAC AGTCACTGAG AGCGACATCC GTACGGAGGA  
 TTCTTTGGG GTTACCCAA GAGCATACTA TGGCGACGA AACTGAGGTG TCAGTGAATC TCGCTGTAGG CATGCCCTCCT

---

+2 A I Y Q C C D L D P Q A R V A I K S L T E R L Y V G  
 6241 GGCAATCTAC CAATGTTGTG ACCTCGACCC CCAAGCCCGC GTGGCCATCA AGTCCCTCAC CGAGAGGCTT TATGTTGGGG  
 CCGTTAGATG GTTACAACAC TGGAGCTGGG GGTCGGCG CACCGTAGT TCAGGGAGTG GCTCTCCGAA ATACAACCCC

---

+2 G P L T N S R G E N C G Y R R C R A S G V L T T S C G  
 6321 GCCCTCTTAC CAATTCAAGG GGGGAGAACT GCGGCTATCG CAGGTGCCGC GCGAGCGCG TACTGACAAC TAGCTGTGGT  
 CGGGAGAATG GTTAAGTTCC CCCCTTTGA CGCCGATAGC GTCCACGGCG CGCTCGCCGC ATGACTGTTG ATCGACACCA

---

+2 N T L T C Y I K A R A A C R A A G L Q D C T M L V C G  
 6401 AACACCTCA CTTGCTACAT CAAGGCCCGG GCAGCCTGTC GAGCCGAGG GCTCCAGGAC TGCACCATGC TCGTGTGTGG  
 TTGTGGAGT GAACGATGTA GTTCCGGGCC CGTCGGACAG CTCGGCTGCC CGAGGCTCTG ACGTGGTACG AGCACACACC

---

+2 D D L V V I C E S A G V Q E D A A S L R A F T E A M  
 6481 CGACGACTTA GTCGTTATCT GTGAAAGCGC GGGGGTCCAG GAGGACGCGG CGAGCCTGAG AGCCTCACG GAGGCTATGA  
 GCTGCTGAAT CAGCAATAGA CACTTCGCG CCCCGAGGTC CTCCCTGCC GCTCGGACTC TCGGAAGTGC CTCCGATACT

---

+2 T R Y S A P P G D P P Q P E Y D L E L I T S C S S N V  
 6561 CCAGGTACTC CGCCCCCCCCCT GGGGACCCCC CACAACCAGA ATACGACTTG GAGCTCATAA CATCATGCTC CTCCAACGTG  
 GGTCCATGAG GCGGGGGGGG CCCCTGGGG GTGTTGGTCT TATGCTGAAC CTCGAGTATT GTAGTACGAG GAGGTTGCAC

---

+2 S V A H D G A G K R V Y Y L T R D P T T P L A R A A W  
 6641 TCAGTCGCC ACGACGGCGC TGGAAAGAGG GTCTACTACC TCACCCGTGA CCCTACAAACC CCCCTCGCGA GAGCTGCCTG  
 AGTCAGCGGG TGCTGCCCGC ACCTTCTCC CAGATGATGG AGTGGCACT GGGATGTTGG GGGGAGCGCT CTCGACGCAC

---

+2 E T A R H T P V N S W L G N I I M F A P T L W A R M  
 6721 GGAGACAGCA AGACACACTC CAGTCAATTCTGCTAGGC AACATAATCA TGTGTCGCCCC CACACTGTGG GCGAGGATGA  
 CCTCTGTCGT TCTGTGTGAG GTCAGTTAAG GACCGATCCG TTGTATTAGT ACAACGGGG GTGTGACACC CGCTCCTACT

---

+2 I L M T H F F S V L I A R D Q L E Q A L D C E I Y G A  
 6801 TACTGATGAC CCATTCTTT AGCGTCTTA TAGCCAGGG CCAGCTTGA CAGGCCCTCG ATTGCGAGAT CTACGGGGCC  
 ATGACTACTG GGTAAAGAAA TCGCAGGAAT ATCGGTCCCT GTGCAACTT GTCCGGGAGC TAACGCTCTA GATGCCCGG

---

+2 C Y S I E P L D L P P I I Q R L H G L S A F S L H S Y  
 6881 TGCTACTCCA TAGAACCAACT GGATCTACCT CCAATCATTC AAAGACTCCA TGGCCTCAGC GCATTTCAC TCCACAGTTA  
 ACGATGAGGT ATCTGGTGA CCTAGATGGA GTTAGTAAAG TTTCTGAGGT ACCGGAGTCG CGTAAAAGTG AGGTGTCAT

---

+2 S P G E I N R V A A C L R K L G V P P L R A W R H R  
 6961 CTCTCCAGGT GAAATCAATA GGGTGGCGC ATGCCCTAGA AAACCTGGGG TACCGCCCTT GCGAGCTGG AGACACCGGG  
 GAGAGGTCCA CTTAGTTAT CCCACCGCGC TACGGAGTC TTTGAACCC ATGGCGGGAA CGCTCGAACCC TCTGTGGCCC

---

+2 A R S V R A R L L A R G G R A A I C G K Y L F N W A V  
 7041 CCCGGAGCGT CGCGCTAGG CTTCTGGCCA GAGGAGGCAG GGCTGCCATA TGTGGCAAGT ACCTCTTCAA CTGGCAGTA  
 GGGCTCGCA GGCGCGATCC GAAGACCGGT CTCCTCCGTC CCGACGGTAT ACACCGTTCA TGGAGAAGTT GACCCGTCAT

## FIGURE 3 - Page 7

+2 R T K L K L T P I A A A G Q L D L S G W F T A G Y S G  
 7121 AGAACAAAGC TCAAACCTCAC TCCAATAGCG GCCGCTGGCC AGCTGGACTT GTCCGGCTGG TTCACGGCTG GCTACAGCGG  
 TCTTGTTCG AGTTGAGTG AGGTTATCGC CGGCGACCGG TCGACCTGAA CAGGCCGACC AAGTGCCGAC CGATGTGCC

+2 G D I Y H S V S H A R P R W I W F C L L L L A A G V  
 7201 GGGAGACATT TATCACAGCG TGTCTCATGC CGGGCCCCGC TGGATCTGGT TTTGCCTACT CCTGCTTGCT GCAGGGGTAG  
 CCCTCTGTAA ATAGTGTGCG ACAGAGTACG GGCCGGGGCG ACCTAGACCA AAACGGATGA GGACGAACGA CGTCCCCATC

+2 G I Y L L P N R  
 7281 GCATCTACCT CCTCCCCAAC CGATGAAGGT TGGGGTAAAC ACTCCGGCCT AAAAAAAA AAAAATCTAG AAAGGCGCGC  
 CGTAGATGGA GGAGGGGTTG GCTACTCCA ACCCCATTG TGAGGCCGA TTTTTTTTT TTTTAGATC TTTCCGCGC

BamHI MluI

7361 CAAGATATCA AGGATCCACT ACGCGTTAGA GCTCGCTGAT CAGCCTCGAC TGTGCCTTCT AGTTGCCAGC CATCTGTGTTG  
 GTTCTATAGT TCCTAGGTGA TGCGCAATCT CGAGCGACTA GTCGGAGCTG ACACGGAAGA TCAACGGTCG GTAGACAACA

7441 TTGCCCCCTCC CCCGTGCCTT CCTTGACCCCT GGAAGGTGCC ACTCCCACGT TCCTTCCTA ATAAAATGAG GAAATTGCAT  
 AACGGGGAGG GGGCACGGAA GGAACGGGA CCTTCCACGG TGAGGGTGAC AGGAAAGGAT TATTTACTC CTTAACGTA

7521 CGCATTGTCT GAGTAGGTGT CATTCTATTG TGGGGGGTGG GGTGGGGCAG GACAGCAAGG GGGAGGATTG GGAAGACAAT  
 GCGTAACAGA CTCATCCACA GTAAGATAAG ACCCCCCACC CCACCCCGTC CTGTCGTTCC CCCTCTTAAC CCTTCTGTTA

7601 ACCAGGCATG CTGGGGGAGCT CTTCCGCTTC CTCGCTCACT GACTCGCTGC GCTCGTGT CGTGGCTGCC CGAGCGGTAT  
 TCGTCCGTAC GACCCCTCGA GAAGGCGAAG GAGCGAGTGA CTGAGCGACG CGAGCCAGCA AGCCGACGCC GCTGCCATA

7681 CAGCTCACTC AAAGGCGGTA ATACGGTTAT CCACAGAAC AGGGGATAAC GCAGGAAAGA ACATGTGAGC AAAAGGCCAG  
 GTCGAGTGAG TTTCCGCCAT TATGCCAATA GGTGTCTTAG TCCCCTATTG CGTCCTTCT TGTACACTCG TTTCCGGTC

7761 CAAAAGGCCA GGAACCGTAA AAAGGCCGCG TTGCTGGCGT TTTCCATAG GCTCCGGCCC CCTGACGAGC ATCACAAAAA  
 GTTTCCGGT CCTTGGCATT TTTCCGGCGC AACGACCGCA AAAAGGTATC CGAGGCCGGG GGACTGCTCG TAGTGTTTT

7841 TCGACGCTCA AGTCAGAGGT GGCGAAACCC GACAGGACTA TAAAGATACC AGGCGTTCC CCCTGGAAGC TCCCTCGTGC  
 AGCTGCGAGT TCAGTCTCCA CCGCTTGGG CTGTCCTGAT ATTTCTATGG TCCGCAAAGG GGGACCTTCG AGGGAGCACG

7921 GCTCTCTGT TCCGACCCCTG CCGCTTACCG GATACTGTC CGCCTTCTC CCTTCGGAA GCCTGGCGCT TTCTCAATGC  
 CGAGAGGACA AGGCTGGGAC GGCAGATGGC CTATGGACAG GCGGAAAGAG GGAAGCCCTT CGCACCGCGA AAGAGTTACG

8001 TCACCGCTGTA GGTATCTAG TTCGGTGTAG GTCTGTCGCT CCAAGCTGGG CTGTGTGCAC GAACCCCCCG TTCAGCCGA  
 AGTGCACAT CCATAGAGTC AAGCCACATC CAGCAAGCGA GGTCGACCC GACACACGT CTGGGGGGC AAGTCGGCCT

8081 CCGCTGCGCC TTATCCGGTA ACTATCGTCT TGAGTCCAAC CCGGTAAGAC ACGACTTATC GCCACTGGCA GCAGCCACTG  
 GGCGACGCCG AATAGGCCAT TGATAGCAGA ACTCAGGTTG GGCCATTCTG TGCTGAATAG CGGTGACCGT CGTCGGTGC

8161 GTAACAGGAT TAGCAGAGCG AGGTATGTAG GCGGTGCTAC AGAGTTCTTG AAGTGGTGGC CTAACACTACGG CTACACTAGA  
 CATTGTCCTA ATCGTCTCGC TCCATACATC CGCCACGATG TCTCAAGAAC TTCACCCACCG GATTGATGCC GATGTGATCT

8241 AGGACAGTAT TTGGTATCTG CGCTCTGCTG AAGCCAGTTA CCTTCGGAAA AAGAGTTGGT AGCTCTTGAT CCGGCAAACA  
 TCCTGTCATA AACCCTAGAC GCGAGACGAC TTGGTCAAT GGAAGCCTT TTCTCAACCA TCGAGAACTA GGCGTTGT

8321 AACCAACCGCT GGTAGCGGTG GTTTTTTGT TTGCAAGCAG CAGATTACGC GCAGAAAAAA AGGATCTAA GAAGATCCTT  
 TTGGTGGCGA CCATGCCAC CAAAAAAACA AACGTTCGTC GTCTAATGCG CGTCTTTTT TCTTAGAGTT CTTCTAGGAA

8401 TGATCTTTC TACGGGGTCT GACGCTCACT GGAAACGAAA CTCACGTTAA GGGATTTGG TCATGAGATT ATCAAAAGG  
 ACTAGAAAAG ATGCCAGA CTGCGAGTCA CCTTGCTTT GAGTGAATT CCCTAAAACC AGTACTCTAA TAGTTTCC

## FIGURE 3 - Page 8

8481 ATCTTCACCT AGATCCTTTT AAATTAAGGGAA TGAAGTTTA AATCAATCTA AAGTATATAT GAGTAAACTT GGTCTGACAG  
. TAGAAGTGGAA TCTAGGAAAA TTTAATTTC ACTTCAAAAT TTAGTTAGAT TTCATATATA CTCATTTGAA CCAGACTGTC

---

8561 TTACCAATGC TTAATCAGTG AGGCACCTAT CTCAGCGATC TGTCTATTTC GTTCATCCAT AGTTGCCTGA CTCCCCGTCG  
AATGGTTACG AATTAGTCAC TCCGTGGATA GAGTCGCTAG ACAGATAAAG CAAGTAGGTA TCAACGGACT GAGGGGCAGC

---

8641 TGTAGATAAC TACGATACGG GAGGGCTTAC CACTCTGGCCC CAGTGCCTGCA ATGATACCGC GAGACCCACG CTCACCGGC  
ACATCTATTG ATGCTATGCC CTCCCGAATG GTAGACCGGG GTCACGACGT TACTATGGCG CTCTGGGTGC GAGTGGCCGA

---

8721 CCAGATTAT CAGCAATAAA CCAGCCAGCC GGAAGGGCCG AGCGCAGAAG TGGTCCCTGCA ACTTTATCCG CCTCCATCCA  
GGTCTAAATA GTCGTTATTG GTCGGTCGG CCTTCCCGC TCGCGTCTC ACCAGGACGT TGAAATAGGC GGAGGTAGGT

---

8801 GTCTATTAAT TGTTGCCGGG AAGCTAGAGT AAGTAGTTCG CCAGTTAATA GTTTGCGCAA CGTTGTTGCC ATTGCTACAG  
CAGATAATTA ACAACGGCCC TTCGATCTCA TTCACTCAAGC GGTCAATTAT CAAACCGCTT GCAACAAACGG TAACGATGTC

---

8881 GCATCGTGGT GTCACGCTCG TCGTTGGTA TGGCTTCATT CAGCTCCGGT TCCCAACGAT CAAGGCAGT TACATGATCC  
CGTAGCACCA CAGTGCAGGC AGCAAACCAT ACCGAAGTAA GTCGAGGCCA AGGGTTGCTA GTTCCGCTCA ATGTACTAGG

---

8961 CCCATGTTGT GCAAAAAAGC GGTTAGCTCC TTCCGTCCTC CGATCGTTGT CAGAAGTAAG TTGGCCGCAG TGTTATCACT  
GGGTACAACA CGTTTTTCG CCAATCGAGG AAGCCAGGAG GCTAGCAACA GTCTTCATTCA AACCAGCGTC ACAATAGTG

---

9041 CATGGTTATG GCAGCACTGC ATAATTCTCT TACTGTCTG CCATCCGTA GATGCTTTTC TGTGACTGGT GAGTACTCAA  
GTACCAATAC CGTCGTGAGC TATTAAGAGA ATGACAGTAC GGTAGGCATT CTACGAAAG AACTGACCA CTCATGAGTT

---

9121 CCAAGTCATT CTGAGAATAG TGTATGCCGC GACCGAGTTG CTCTTGCCCG GCGTCAATAC GGGATAATAC CGCGCCACAT  
GGTTCAGTAA GACTCTTATC ACATAACGGC CTGGCTCAAC GAGAACGGGC CGCAGTTATG CCTTATTATG CGCGGGTGT

---

9201 AGCAGAACTT TAAAAGTGCT CATCATGGAA AAACGTTCTT CGGGGCGAAA ACTCTCAAGG ATCTTACCGC TGTTGAGATC  
TCGTCTTGAA ATTTTCACGA GTAGAACCT TTTGCAAGAA GCCCCGCTT TGAGAGTTCC TAGAATGGCG ACAACTCTAG

---

9281 CAGTCGATG TAACCCACTC GTGCACCAA CTGATCTTCA GCATCTTTA CTTTCACCAAG CGTTCTGGG TGAGCAAAA  
GTCAAGCTAC ATTGGGTGAG CACGTGGTT GACTAGAAGT CGTAGAAAAT GAAAGTGGTC GCAAAGACCC ACTCGTTTT

---

9361 CAGGAAGGCA AAATGCCGCA AAAAAGGGAA TAAGGGCGAC ACGGAAATGT TGAATACTCA TACTCTTCCCT TTTTCAATAT  
GTCCTTCCGT TTTACGGCGT TTTTCCCTT ATTCCCGCTG TGCCCTTACA ACTTATGAGT ATGAGAAGGA AAAAGTTATA

---

9441 TATTGAAGCA TTTATCAGGG TTATTGTCTC ATGAGCGGAT ACATATTGA ATGTATTTAG AAAAATAAAC AAATAGGGT  
ATAACTTCGT AAATAGTCCC AATAACAGAG TACTCGCCTA TGTATAAACT TACATAAATC TTTTTATTTG TTTATCCCC

---

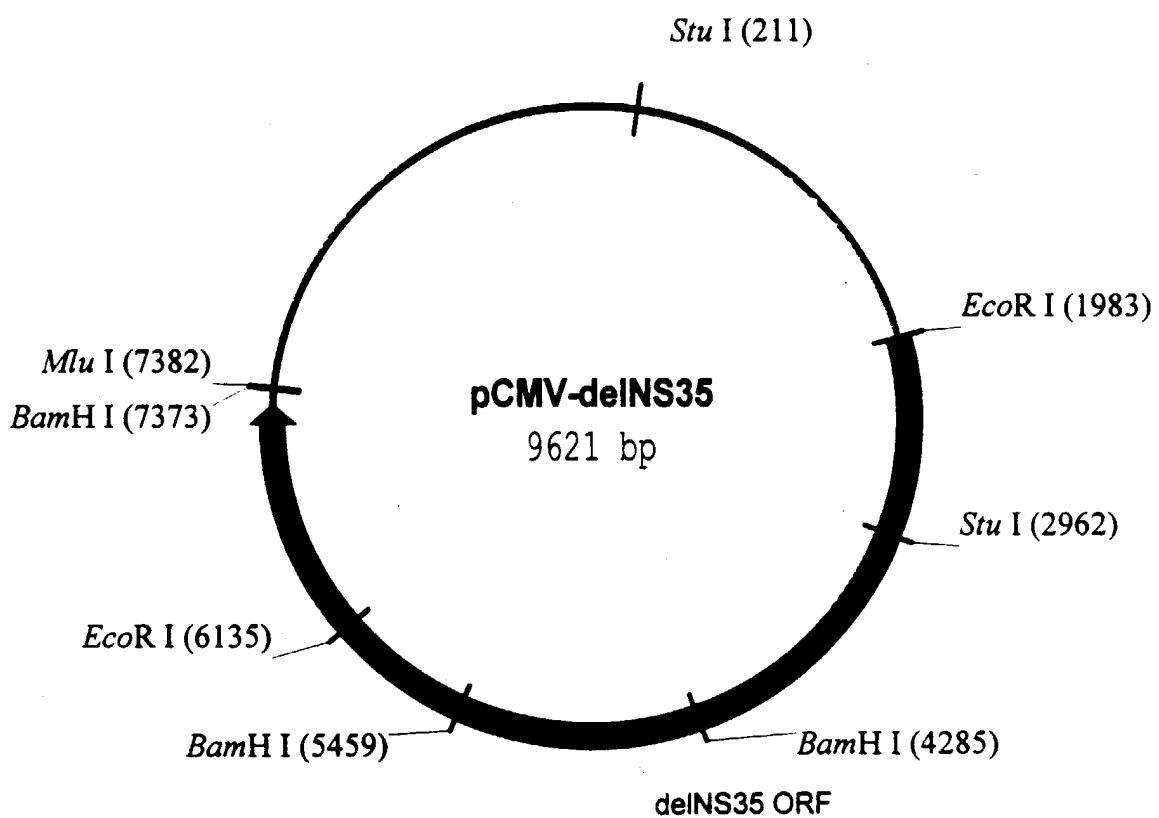
9521 TCCGCGCACA TTTCCCGAA AAGTGCACCC TGACGTCTAA GAAACCATTA TTATCATGAC ATTAACCTAT AAAAATAAGGC  
AGGCGCGTGT AAAGGGGCTT TTCACGGTGG ACTGCAGATT CTTGGTAAT AATAGTACTG TAATTGGATA TTTTTATCCG

---

9601 GTATCACGAG GCCCTTTCGT C  
CATAGTGCTC CGGGAAAGCA G

---

**FIGURE 4**



## FIGURE 5 - Page 1

1 TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG GAGACGGTCA CAGCTTGTCT GTAAGCGGAT  
AGCGCGAAA GCCACTACTG CCACTTTGG AGACTGTGTA CGTCGAGGGC CTCTGCCAGT GTCGAACAGA CATTGCCCTA

81 GCCGGGAGCA GACAAGCCCG TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG CGGCATCAGA  
CGGCCCTCGT CTGTTCGGC AGTCCCAGC AGTCGCCAC AACCGCCCAC AGCCCCGACC GAATTGATAC GCCGTAGTCT

StuI  
-----

161 GCAGATTGTA CTGAGAGTGC ACCATATGAA GCTTTTGCA AAAGCCTAGG CCTCCAAAAA AGCCTCCTCA CTACTTCTGG  
CGTCTAACAT GACTCTCAGG TGGTATACTT CGAAAAACGT TTTCGGATCC GGAGGTTTT TCGGAGGAGT GATGAAGACC

241 AATAGCTCAG AGGCCGAGGC GGCTCGGCC TCTGCATAAA TAAAAAAAAT TAGTCAGCCA TGGGGCGGAG AATGGCGGA  
TTATCGAGTC TCCGGCTCCG CCGGAGCCGG AGACGTATTT ATTTTTTTA ATCAGTCGGT ACCCCGCCCTC TTACCCGCCT

321 ACTGGGCAGGG GAGGGAATTA TTGGCTATTG GCCATTGCAT ACGTTGTATC TATATCATAA TATGTACATT TATATTGGCT  
TGACCCGCCCTC CTCCTTAAT AACCGATAAC CGGTAAACGTA TGCAACATAG ATATAGTATT ATACATGTAATATAACCGA

401 CATGTCCAAT ATGACCGCCA TGTTGACATT GATTATTGAC TAGTTATTA TAGTAATCAA TTACGGGTC ATTAGTTCAT  
GTACAGGTAA TACTGGCGGT ACAACTGTAA CTAATAACTG ATCAATAATT ATCATTAGTT AATGCCCAAG TAATCAAGTA

481 AGCCCATAATA TGAGGTTCCG CGTTACATAA CTTACGGTAA ATGGCCGCC TGGCTGACCG CCCAACGACC CCCGCCATT  
TCGGGTATAT ACCTCAAGGC GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC GGGTTGCTGG GGGCGGGTAA

561 GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCATAA GGGACTTCC ATTGACGTCA ATGGGTGGAG TATTTACGGT  
CTGCAGTTAT TACTGCATAC AAGGGTATCA TTGCGGTTAT CCCTGAAAGG TAACTGCAGT TACCCACCTC ATAAATGCCA

641 AAACGTCCCA CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTCCGCC CCTATTGACG TCAATGACGG TAAATGGGCC  
TTTGACGGGT GAACCGTCAT GTAGTTACA TAGTATACGG TTCAGGGCGG GGATAACTGC AGTTACTGCC ATTTACCGGG

721 GCCTGGCATT ATGCCAGTA CATGACCTTA CGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC  
CGGACCGTAA TACGGGTCA GTACTGGAT GCCCTGAAAG GATGAACCGT CATGTAGATG CATAATCAGT AGCGATAATG

801 CATGGTGATG CGGTTTGGC AGTACACCAA TGGCGTGGA TAGCGGTTTG ACTCACGGG ATTTCCAAGT CTCCACCCCA  
GTACCACTAC GCCAAACCG TCATGTGGTT ACCCGCACCT ATCGCCAAAC TGAGTGCCTC TAAAGGTCA GAGGTGGGT

881 TTGACGTCAA TGGGAGTTG TTTGGCACC AAAATCAACG GGACTTTCCA AAATGTCGTAA TAAACCCCGC CCCGTTGACG  
AACTGCAGTT ACCCTCAAAC AAAACCGTGG TTTAGTTGC CCTGAAAGGT TTTACAGCAT TATTGGGGCG GGGCAACTGC

961 CAAATGGCG GTAGGCGTGT ACGGTGGAG GTCTATATAA GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG  
GTTTACCCGC CATCCGCACA TGCCACCTC CAGATATATT CGTCTCGAGC AAATCACTTG GCAGTCTAGC GGACCTCTGC

1041 CCATCCACGC TGTGTTGACC TCCATAGAAAG ACACCGGGAC CGATCCAGCC TCCGCGGCCG GGAACGGTGC ATTGGAACGC  
GGTAGGTGCG ACAAAACTGG AGGTATCTTC TGTGGCCCTG GCTAGGTGG AGGCGCCGGC CCTTGCCACG TAACCTTGCG

1121 GGATTCCCCG TGCCAAGAGT GACGTAAGTA CGCCTATAG ACTCTATAGG CACACCCCTT TGGCTTCTAT GCATGCTATA  
CCTAAGGGC ACGTTCTCA CTGCATTCA GGCAGATATC TGAGATATCC GTGTGGGAA ACCGAGAATA CGTACGATAT

1201 CTGTTTTGG CTTGGGGCCT ATACACCCCG GCTCCTTATG CTATAGGTGA TGGTATAGCT TAGCCTATAG GTGTGGGTAA  
GACAAAAACC GAACCCCGGA TATGTGGGG CGAGGAATAC GATATCCACT ACCATATCGA ATCGGATATC CACACCCAT

1281 TTGACCATTAA TTGACCACTC CCCTATTGGT GACGATACTT TCCATTACTA ATCCATAACA TGGCTCTTG CCACAACTAT  
AACTGGTAAT AACTGGTGAG GGGATAACCA CTGCTATGAA AGGTAATGAT TAGGTATTGT ACCGAGAAC GGTGTTGATA

1361 CTCTATTGGC TATATGCCAA TACTCTGTC TTCAGAGACT GACACGGACT CTGTATTTT ACAGGATGGG GTCCATTAT  
GAGATAACCG ATATACGGTT ATGAGACAGG AAGTCTGTA CTGTGCCTGA GACATAAAA TGCTACCC CAGGTAAATA

## FIGURE 5 - Page 2

1441 TATTTACAAA TTCACATATA CAACAACGCC GTCCCCGTG CCCGCAGTTT TTATTAACAA TAGCGTGGGA TCTCCGACAT  
ATAAAATGTTT AAGTGTATAT GTTGGTGCAG CAGGGGCAC GGGCGTCAA AATAATTGT ATCGCACCT AGAGGCTGTA

1521 CTCGGTAGC TTGTCGGAC ATGGGCTCTT CTCCGGTAGC GGCGGAGCTT CCACATCCGA GCCCTGGTCC CATCCGTCCA  
GAGCCCATGC ACAAGGCCTG TACCCGAGAA GAGGCCATCG CGGCCTCGAA GGTGTAGGCT CGGGACCAGG GTAGGCAGGT

1601 GCGGCTCATG GTGCGCTCGGC AGCTCCTTGC TCCTAACAGT GGAGGCCAGA CTTAGGCACA GCACAATGCC CACCACC  
CGCCGAGTAC CAGCGAGCCG TCGAGGAACG AGGATTGTCA CCTCCGGTCT GAATCCGTGT CGTGTACGG GTGGTGGTGG

1681 AGTGTGCCGC ACAAGGCCGT GGCGGTAGGG TATGTGTCTG AAAATGAGCT CGGAGATTGG GCTCGCACCT GGACGCCAGAT  
TCACACGGCG TGTTCGGCA CGGCCATCCC ATACACAGAC TTTTACTCGA GCCTCTAACCG CAGCGTGGAA CCTCGCTCTA

1761 GGAAGACTTA AGGCAGCGGC AGAAGAAGAT GCAGGCAGCT GAGTTGTTGT ATTCTGATAA GAGTCAGAGG TAACTCCCGT  
CCTTCTGAAT TCCGTCGGC TCTTCTCTA CGTCCGTGCA CTCAACAACA TAAGACTATT CTCAGTCTCC ATTGAGGGCA

1841 TGCGGTGCTG TTAACGGTGG AGGGCAGTGT AGTCTGAGCA GTACTCGTTG CTGCCGCGCG CGCCACCAGA CATAATAGCT  
ACGCCACGAC AATTGCCACC TCCCCTCACAC TCAGACTCGT CATGAGCAAC GACGGCGCGC GCGGTGGTCT GTATTATCGA

+2

M A A

EcoRI

1921 GACAGACTAA CAGACTGTTG CTTTCCATGG GTCTTTCTG CAGTCACCGT CGTCGACCTA AGAATTCAAC ATGGCTGCAT  
CTGTCTGATT GTCTGACAAG GAAAGGTACC CAGAAAAGAC GTCACTGGAT TCTTAAGTGG TACCGACGTA

+2 Y A A Q G Y K V L V L N P S V A A T L G F G A Y M S K  
2001 ATGCAGCTCA GGGCTATAAG GTGCTAGTAC TCAACCCCTC TGTTGCTGCA ACACCTGGCT TTGGTGCTTA CATGCTCAAG  
TACGTCGAGT CCCGATATTG CACGATCATG AGTGGGGAG ACAACGACGT TGTGACCCGA AACCACGAAT GTACAGGTTG

+2 A H G I D P N I R T G V R T I T T G S P I T Y S T Y G  
2081 GCTCATGGGA TCGATCCTAA CATCAGGACC GGGGTGAGAA CAATTACCAAC TGGCAGCCCC ATCACGTACT CCACCTACGG  
CGAGTACCCCT AGCTAGGATT GTAGTCTGG CCCCACTCTT GTTAATGGTG ACCGTCGGGG TAGTGCATGA GGTGGATGCC

+2 K F L A D G G C S G G A Y D I I I C D E C H S T D A  
2161 CAAGTTCTT GCCGACGGCG GGTGCTCGGG GGGCGCTTAT GACATAATAA TTTGTGACGA GTGCCACTCC ACGGATGCCA  
GTTCAAGGAA CGGCTGCCGC CCACGAGCCC CCCCGAATA CTGTATTATT AAACACTGCT CACGGTGAGG TGCCTACGGT

+2 T S I L G I G T V L D Q A E T A G A R L V V L A T A T  
2241 CATCCATCTT GGGCATTGGC ACTGTCTTG ACCAACGAGA GACTGCGGGG GCGAGACTGG TTGTGCTCGC CACCGCCACC  
GTAGGTAGAA CCCGTAACCG TGACAGGAAC TGGTCTGCT CTGACGCCCG CGCTCTGACC AACACGAGCG GTGGCGGTGG

+2 P P G S V T V P H P N I E E V A L S T T G E I P F Y G  
2321 CCTCCGGGCT CCGTCACTGT GCCCCATCCC AACATCGAGG AGGTTGCTCT GTCCACCACC GGAGAGATCC CTTTTACGG  
GGAGGCCGA GGCAGTGACA CGGGGTAGGG TTGTAGCTCC TCCAACGAGA CAGGTGGTGG CCTCTCTAGG GAAAAATGCC

+2 K A I P L E V I K G G R H L I F C H S K K K C D E L  
2401 CAAGGCTATC CCCCTCGAAG TAATCAAGGG GGGGAGACAT CTACATCTCT GTCAATTCAA GAAGAAGTGC GACGAACCTG  
GTTCCGATAG GGGGAGCTTC ATTAGTCTCC CCCCTCTGTA GAGTAGAAGA CAGTAAGTTT CTTCTTCACG CTGCTTGAGC

+2 A A K L V A L G I N A V A Y Y R G L D V S V I P T S G  
2481 CCGCAAAGCT GGTGCGATTG GGCATCAATG CCGTGGCCTA CTACCGCGGT CTTGACGTGT CCGTCATCCC GACCAGCGC  
GGCGTTCGA CCAGCGTAAC CGTAGTTAC GGCACCGGAT GATGGCGCA GAACCTGACA GGCAGTAGGG CTGGTCGCGC

+2 D V V V V A T D A L M T G Y T G D F D S V I D C N T C  
2561 GATGTTGTCG TCGTGGCAAC CGATGCCCTC ATGACCGGCT ATACCGGCA CTTGACTCG GTGATAGACT GCAATACGTG  
CTACAACAGC AGCACCGTTG GCTACGGGAG TACTGGCGA TATGGCCGT GAAGCTGAGC CACTATCTGA CGTTATGCAC

## FIGURE 5 - Page 3

+2 V T Q T V D F S L D P T F T I E T I T L P Q D A V S  
 2641 TGTCACCCAG ACAGTCGATT TCAGCCTGA CCCTACCTTC ACCATTGAGA CAATCACGCT CCCCAAGAT GCTGTCTCCC  
 ACAGTGGTC TGTCAGCTAA AGTCGGAAC GGGATGGAAG TGGTAACTCT GTTAGTGCAGA GGGGGTTCTA CGACAGAGGG

---

+2 R T Q R R G R T G R G K P G I Y R F V A P G E R P S G  
 2721 GCACTCAACG TCGGGGCAGG ACTGGCAGGG GGAAGCCAGG CATCTACAGA TTTGTGGCAC CGGGGGAGCG CCCCCTCCGGC  
 CGTGAGTTGC AGCCCCGTCC TGACCGTCCC CCTTCGGTCC GTAGATGTCT AAACACCGTG GCCCCCTCGC GGGGAGGGCG

---

+2 M F D S S V L C E C Y D A G C A W Y E L T P A E T T V  
 2801 ATGTCGACT CGTCCGTCT CTGTGAGTGC TATGACGCAG GCTGTGCTTG GTATGAGCTC ACGCCCGCCG AGACTACAGT  
 TACAAGCTGA GCAGGCAGGA GACACTCACG ATACTGCGTC CGACACGAAC CATACTCGAG TGCGGGCGGC TCTGATGTCA

---

+2 R L R A Y M N T P G L P V C Q D H L E F W E G V F T  
 StuI  
 ~~~

2881 TAGGCTACGA GCGTACATGA ACACCCCCGGG GCTTCCCGTG TGCCAGGACC ATCTTGAATT TTGGGAGGGC GTCTTACAG  
 ATCCGATGCT CGCATGTACT TGTGGGGCC CGAAGGGCAC ACGGTCTGG TAGAACTTAA AACCCCTCCCG CAGAAATGTC

---

+2 G L T H I D A H F L S Q T K Q S G E N L P Y L V A Y Q  
 StuI  
 ~~~

2961 GCCTCACTCA TATAGATGCC CACTTTCTAT CCCAGACAAA GCAGAGTGGG GAGAACCTTC CTTACCTGGT AGCGTACCAA  
 CGGAGTGAGT ATATCTACGG GTGAAAGATA GGGTCTGTT CGTCTCACCC CTCTTGAAG GAATGGACCA TCGCATGGT

---

+2 A T V C A R A Q A P P P S W D Q M W K C L I R L K P T  
 3041 GCCACCGTGT GCGCTAGGGC TCAAGCCCT CCCCATCGT GGGACCAGAT GTGGAAGTGT TTGATTGCGC TCAAGCCAC  
 CGGTGGCACA CGCGATCCCG AGTCGGGGA GGGGTAGCA CCCTGGTCTA CACCTTCACA AACTAAGCGG AGTTCGGGTG

---

+2 L H G P T P L L Y R L G A V Q N E I T L T H P V T K  
 3121 CCTCCATGGG CCAACACCCC TGCTATACAG ACTGGCGCT GTTCAGAATG AAATCACCC GACGCACCCA GTCACCAAAT  
 GGAGGTACCC GGTTGTGGGG ACGATATGTC TGACCCGCGA CAAGTCTTAC TTTAGTGGGA CTGCGTGGGT CAGTGGTTA

---

+2 Y I M T C M S A D L E V V T S T W V L V G G V L A A L  
 3201 ACATCATGAC ATGCATGTCG GCGACCTGG AGGTCGTCAC GAGCACCTGG GTGCTCGTT GCAGCGTCCT GGCTGCTTIG  
 TGTAGTACTG TAGTACAGC CGGCTGGACC TCCAGCAGTG CTCGTGGACC CACGAGCAAC CGCCGCGAGGA CGGACGAAAC

---

+2 A A Y C L S T G C V V I V G R V V L S G K P A I I P D  
 3281 GCCCGTATT GCCTGTCAAC AGGCTCGTG GTCATAGTGG GCAGGGCTGT CTTGTCCGGG AAGCCGGCAA TCATACCTGA  
 CGGCGATAA CGGACAGTTG TCCGACGAC CAGTATCACC CGTCCCAGCA GAACAGGCC TTCGGCCGT AGTATGGACT

---

+2 R E V L Y R E F D E M E E C S Q H L P Y I E Q G M M  
 3361 CAGGGAAGTC CTCTACCGAG AGTCGATGA GATGGAAGAG TGCTCTCAGC ACTTACCGTA CATCGAGCAA GGATGATGC  
 GTCCCTCAG GAGATGGCTC TCAAGCTACT CTACCTCTC ACGAGAGTCG TGAATGGCAT GTAGCTCGTT CCCTACTACG

---

+2 L A E Q F K Q K A L G L L Q T A S R Q A E V I A P A V  
 3441 TCGCCGAGCA GTTCAGCGAG AAGGCCCTCG GCCTCTCGA GACCGCGTCC CGTCAGGCAG AGGTTATCGC CCCTGCTGTC  
 AGCGGCTCGT CAAGTTCGTC TTCCGGGAGC CGGAGGACGT CTGGCGCAGG GCAGTCCGTC TCCAATAGCG GGGACGACAG

---

+2 Q T N W Q K L E T F W A K H M W N F I S G I Q Y L A G  
 3521 CAGACCAACT GGCAAAACT CGAGACCTTC TGGCGAAGC ATATGTGGAA CTTCATCAGT GGGATACAAT ACTTGGCGGG  
 GTCTGGTTGA CGGTTTTGA GCTCTGGAA ACCCGCTTCG TATACACCTT GAAGTAGTCA CCCTATGTTA TGAACCGCCC

---

+2 L S T L P G N P A I A S L M A F T A A V T S P L T T  
 3601 CTTGTCAACG CTGCGCTGGTA ACCCCGCCAT TGCTCATTG ATGGCTTTA CAGCTCGTGT CACCAAGCCCA CTAACCAACTA  
 GAACAGTTGC GACGGACCAT TGGGGCGTA ACGAAGTAAC TACCGAAAAT GTCGACGACA GTGGTCGGGT GATTGGTGAT

---

+2 S Q T L L F N I L G G W V A A Q L A A P G A A T A F V  
 3681 GCCAAACCCCT CCTCTTCAC ATATTGGGGG GGTGGGTGGC TGCCCGCTC GCGCCCCCG GTGCCGCTAC TGCCCTTGTG  
 CGGTTGGGA GGAGAAGTTG TATAACCCCC CCACCCACCG ACGGGTCGAG CGGGGGGGC CACGGCGATG ACGGAAACAC

## FIGURE 5 - Page 4

+2 G A G L A G A A I G . S V G L G K V L I D I L A G Y G A  
 3761. GGCCTGGCT TAGCTGGCGC CGCCATCGGC AGTGTGGAC TGGGGAAAGT CCTCATAGAC ATCCTTGAG GGTATGGCGC  
 CCGCGACCGA ATCGACCGCG GCGGTAGCCG TCACAACCTG ACCCCTTCCA GGAGTATCTG TAGGAACGTC CCATACCGCG

---

+2 G V A G A L V A F K I M S G E V P S T E D L V N L L  
 3841 GGGCTGGCG GGAGCTCTG TGGCATCAA GATCATGAGC GGTGAGGTCC CCTCCACGGA GGACCTGGTC AATCTACTGC  
 CCCGCACCGC CCTCGAGAAC ACCGTAAGTT CTAGTACTCG CCACTCCAGG GGAGGTGCCT CCTGGACCAAG TTAGATGACG

---

+2 P A I L S P G A L V V G V V C A A I L R R H V G P G E  
 3921 CCGCCATCCT CTCGCCCGA GCCCTCGTAG TCGGCTGGT CTGTGCAGCA ATACTGCAGCC GGACGTTGG CCCGGCGAG  
 GGCGTAGGA GAGCGGGCCT CGGGAGCAGC AGCCGACCA GACACGTCGT TATGACGCGG CCGTGCAACC GGGCCCGCTC

---

+2 G A V Q W M N R L I A F A S R G N H V S P T H Y V P E  
 4001 GGGGCAGTGC AGTGGATGAA CCGCTGATA GCCTTCGCT CCCGGGGAA CCATGTTCC CCCACGCACT ACGTGCAGGA  
 CCCCGTCACG TCACCTACTT GGCGACTAT CGGAAGCGGA GGGCCCCCTT GGTACAAAGG GGGTGCAGTGA TGCACGGCCT

---

+2 S D A A A R V T A I L S S L T V T Q L L R R L H Q W  
 4081 GAGCGATGCA GCTGCCCGCG TCACTGCCAT ACTCAGCAGC CTCACTGTAA CCCAGCTCCT GAGGCGACTG CACCACTGGGA  
 CTCGCTACGT CGACGGCGC AGTGCAGGT TGAGTCGTG GAGTGACATT GGGTCGAGGA CTCGCTGAC GTGGTACACT

---

+2 I S S E C T T P C S G S W L R D I W D W I C E V L S D  
 4161 TAAGCTCGGA GTGTACCACT CCATGCTCCG GTTCTGGCT AAGGGACATC TGGGACTGGA TATGCGAGGT GTTGAGCGAC  
 ATTGAGCCT CACATGGTGA GGTACGAGGC CAAGGACCGA TTCCCTGTAG ACCCTGACCT ATACGCTCCA CAACTCGCTG

---

+2 F K T W L K A K L M P Q L P G I P F V S C Q R G Y K G  
 ~~~~~~  
 4241 TTTAAGACCT GGCTAAAAGC TAAGCTCATG CCACAGCTGC CTGGGATCCC CTTTGTGTCC TGCCAGCGCG GGTATAAGGG  
 AAATTCTGGA CCGATTTTCG ATTCGAGTAC GGTGTCGACG GACCCTAGGG GAAACACAGG ACGGTGCGCG CCATATTCCC

---

+2 V W R G D G I M H T R C H C G A E I T G H V K N G T  
 4321 GGTCTGGCGA GGGGACGGCA TCATGCACAC TCGTGCACAC TGTGGAGCTG AGATCACTGG ACATGTCAA AACGGGACGA  
 CCAGACCGCT CCCCTGCCGT AGTACGTGT AGCGACGGT ACACCTCGAC TCTAGTGACC TGTACAGTT TTGCCCTGCT

---

+2 M R I V G P R T C R N M W S G T F P I N A Y T T G P C  
 4401 TGAGGATCGT CGGTCTAGG ACCTGCAGGA ACATGTGGAG TGGGACCTTC CCCATTAATG CCTACACCAC GGGCCCCGT  
 ACTCCTAGCA GCCAGGATCC TGGACGTCTC ACCCTGAAAG GGGTAATTAC GGATGTGGTG CCCGGGGACA

---

+2 T P L P A P N Y T F A L W R V S A E E Y V E I R Q V G  
 4481 ACCCCCCCTTC CTGCGCCGAA CTACACGTT GCGCTATGGA GGGTGTCTGC AGAGGAATAC GTGGAGATAA GGCAGGTGGG  
 TGGGGGAAG GACGCGGCTT GATGTGCAAG CGCGATACCT CCCACAGACG TCTCCTTATG CACCTCTATT CGTCCACCC

---

+2 D F H Y V T G M T T D N L K C P C Q V P S P E F F T  
 4561 GGACTTCCAC TACGTGACGG GTATGACTAC TGACAATCTT AAATGCCGT GCCAGGTCCC ATCGCCCGAA TTTTCACAG  
 CCTGAAGGTG ATGCACTGCC CATACTGATG ACTGTTAGAA TTTACGGCA CGGTCCAGGG TAGCGGGCTT AAAAAGTGTG

---

+2 E L D G V R L H R F A P P C K P L L R E E V S F R V G  
 4641 AATTGGACGG GGTGCGCTA CATAGTTTG CGCCCCCTG CAAGCCCTG CTGGGGAGG AGGTATCATT CAGAGTAGGA  
 TTAACCTGCC CCACGCGGAT GTATCCAAAC GCGGGGGAC GTTCGGGAAC GACGCCCTCC TCCATAGTAA GTCTCATCCT

---

+2 L H E Y P V G S Q L P C E P E P D V A V L T S M L T D  
 4721 CTCCACGAAT ACCCGTAGG GTCGCAATTG CCTTGCAGC CCGAACCGGA CGTGGCCGTG TTGACGTCCA TGCTCACTGA  
 GAGGTGCTTA TGGGCCATCC CAGCGTTAAT GGAACGCTCG GGCTTGGCT GCACCGGCAC AACTGCAGGT ACGAGTGACT

---

+2 P S H I T A E A A G R R L A R G S P P S V A S S S A  
 4801 TCCCTCCCAT ATAACAGCAG AGGCAGCGG GCGAAGGTG GCGAGGGAT CACCCCCCTC TGTGGCCAGC TCCTCGGCTA  
 AGGGAGGGTA TATTGTCGTC TCCGCCGGCC CGCTCCAAAC CGCTCCCCCTA GTGGGGGAG ACACCGGTG AGGAGGCCAT

## FIGURE 5 - Page 5

+2 S Q L S A P S L K A T C T A N H D S P D A E L I E A N  
 4881 GCCAGCTATC CGCTCCATCT CTCAGGCAA CTTGCACCGC TAACCATGAC TCCCTGATG CTGAGCTCAT AGAGGCCAAC  
 CGGTCGATAG GCGAGGTAGA GAGTTCCGTT GAACGTGGCG ATTGGTACTG AGGGGACTAC GACTCGAGTA TCTCCGGTTG

+2 L L W R Q E M G G N I T R V E S E N K V V I L D S F D  
 4961 CTCCTATGGA GGCAGGAGAT GGGCGGCAAC ATCACCGGG TTGAGTCAGA AAACAAAGTG GTGATTCTGG ACTCCTTCGA  
 GAGGATACCT CCGCCTCTA CCCGCCGTTG TAGTGGTCCC AACTCAGTCT TTTGTTCAC CACTAAGACC TGAGGAAGCT

+2 P L V A E E D E R E I S V P A E I L R K S R R F A Q  
 5041 TCCGCTGTG GCGGAGGAGG ACGAGCGGA GATCTCGTA CCCGCAGAAA TCCTCGGAA GTCTGGAGA TTCGCCAGG  
 AGGCGAACAC CGCCTCTCC TGCTGCCCT CTAGAGGCAT GGGCGTCTT AGGACGCCCT CAGAGCCTCT AAGCGGGTCC

+2 A L P V W A R P D Y N P P L V E T W K K P D Y E P P V  
 5121 CCCTGCCGT TTGGCGCGG CCGGACTATA ACCCCCCGCT AGTGGAGACG TGGAAAAGC CCGACTACGA ACCACCTGTG  
 GGGACGGGCA AACCCGCGCC GGCCTGATAT TGGGGGGCGA TCACCTCTGC ACCTTTTCG GGCTGATGCT TGGTGGACAC

+2 V H G C P L P P P K S P P V P P P R K K R T V V L T E  
 5201 GTCCATGGCT GCCCGCTTCC ACCTCCAAAG TCCCTCCTG TGCCCTCCCGC TCGGAAGAAG CGGACGGTGG TCCTCACTGA  
 CAGGTACCGA CGGGCGAAGG TGGAGGTTTC AGGGGAGGAC ACGGAGGCCG AGCCTTCTC GCCTGCCACC AGGAGTGACT

+2 S T L S T A L A E L A T R S F G S S S T S G I T G D  
 5281 ATCAACCTA TCTACTGCCT TGGCGAGCT CGCCACCAGA AGCTTGCGA GCTCCTCAAC TTCCGGCATT ACGGCGACA  
 TAGTTGGAT AGATGACGGA ACCGGCTCGA GCGGTGGTCT TCGAAACCGT CGAGGAGTTG AAGGCCGAA TGCCCGCTGT

+2 N T T T S S E P A P S G C P P D S D A E S Y S S M P P  
 5361 ATACGACAAC ATCCTCTGAG CCCGCCCCCTT CTGGCTGCC CCCGACTCC GACGCTGAGT CCTATTCCCT CATGCCCCCC  
 TATGCTGTTG TAGGAGACTC GGGGGGGAA GACCGACGGG GGGGCTGAGG CTGCGACTCA GGATAAGGAG GTACGGGGGG

+2 L E G E P G D P D L S D G S W S T V S S E A N A E D V  
 BamHI  
 ~~~~~~

5441 CTGGAGGGGG AGCCTGGGG A T C C G G A T C T T A G C G A C G G G T C A A C G G T C A A C G G T C A A C G G A A C G  
 G A C C T C C C C C T C G G A C C C C T A G G C T A G A A T C G C T G C C A G T A C C G T G T G C G C T A C C G T C T A C A

+2 V C C S M S Y S W T G A L V T P C A A E E Q K L P I  
 5521 CGTGTGCTGC TCAATGTCTT ACTCTGGAC AGGGCGACTC GTCACCCCGT GCGCCGCGGA AGAACAGAAA CTGCCCATCA  
 GCACACGACG AGTTACAGAA TGAGAACCTG TCCGCGTGAG CAGTGGGCA CGCGCGCCT TCTTGTCTT GACGGTAGT

+2 N A L S N S L L R H H N L V Y S T T S R S A C Q R Q K  
 5601 ATGCACTAAG CAACTCGTTG CTACGTCAAC ACAATTGGT GTATTCCACC ACCTCACGCA GTGCTTGCA AAGGCAGAAC  
 TACGTGATTC GTTGAGCAAC GATGCAGTGG TGTAAACCA CATAAGGTGG TGGAGTGCCT CACGAACGGT TTCCGTCTTC

+2 K V T F D R L Q V L D S H Y Q D V L K E V K A A A S K  
 5681 AAAGTCACAT TTGACAGACT GCAAGTTCTG GACAGCCATT ACCAGGACGT ACTCAAGGAG GTTAAAGCAG CGCGCTCAAA  
 TTTCAGTGTAA AACTGTCTGA CGTTCAAGAC CTGTCGGTAA TGGTCCCTC TGAGTTCCCT CAATTCGTC GCCGCAAGTTT

+2 V K A N L L S V E E A C S L T P P H S A K S K F G Y  
 5761 AGTGAAGGCT AACTTGCTAT CCGTAGAGGA AGCTTGAGC CTGACGCC CACACTCAGC CAAATCCAAG TTTGGTTATG  
 TCACTTCCGA TTGAACGATA GGCATCTCCT TCGAACGTG GACTGCAGGG GTGTGAGTCG GTTTAGGTT AACCAATAAC

+2 G A K D V R C H A R K A V T H I N S V W K D L L E D N  
 5841 GGGCAAAAGA CGTCCGTTGC CATGCCAGAA AGGCCGTAAC CCACATCAAC TCCGTGTTGA AAGACCTCT GGAAGACAAT  
 CCCGTTTCT GCAGGCAACG GTACGGTCTT TCCGGCATTG GGTGTAGTT AGGCACACCT TTCTGGAAGA CCTCTGTTA

+2 V T P I D T T I M A K N E V F C V Q P E K G G R K P A  
 5921 GTAACACCAA TAGACACTAC CATCATGGCT AAGAACGAGG TTTCTGCGT TCAGCCTGAG AAGGGGGGTC GTAAGCCAGC  
 CATTGTGGTT ATCTGTGATG GTAGTACCGA TTCTGCTCC AAAAGACGA AGTCGGACTC TTCCCCCAG CATTGGTCG

## FIGURE 5 - Page 6

+2 R L I V F P D L G V R V C E K M A L Y D V V T K L P  
 6001 TCGTCTCATC GTGTTCCCCG ATCTGGCGT GC CGTGTGC GAAAAGATGG CTTTGTACGA CGTGGTTACA AAGCTCCCT  
 AGCAGAGTAG CACAAGGGGC TAGACCCGCA CGCGCACACG CTTTTCTACC GAAACATGCT GCACCAATGT TTGAGGGGA

---

+2 L A V M G S S Y G F Q Y S P G Q R V E F L V Q A W K S  
 EcoRI  
 ~~~~~~

6081 TGGCCGTGAT GGGAAAGCTCC TACGGATTCC AATACTCACC AGGACAGCGG GTTGAATTCC TCGTCAAGC GTGAAAGTCC  
 ACCGGCACTA CCCTTCGAGG ATGCCTAAGG TTATGAGTGG TCCCTGTGCC CAACTTAAGG AGCACGTTG CACCTTCAGG

---

+2 K K T P M G F S Y D T R C F D S T V T E S D I R T E E  
 6161 AAGAAAACCC CAATGGGTT CTCGTATGAT ACCCGCTGCT TTGACTCCAC AGTCACTGAG AGCGACATCC GTACGGAGGA  
 TTCTTTGGG GTTACCCCAA GAGCATACTA TGGCGACGA AACTGAGGTG TCAGTGACTC TCGCTGTAGG CATGCCCTC

---

+2 A I Y Q C C D L D P Q A R V A I K S L T E R L Y V G  
 6241 GGCAATCTAC CAATGTTGTG ACCTCGACCC CCAAGCCCGC GTGGCCATCA AGTCCCTCAC CGAGAGGCTT TATGTTGGG  
 CGGTTAGATG GTTACAACAC TGGAGCTGGG GGTCGGCG CACCGGTAGT TCAGGGAGTG GCTCTCCGAA ATACAACCCC

---

+2 G P L T N S R G E N C G Y R R C R A S G V L T T S C G  
 6321 GCCCTCTTAC CAATTCAAGG GGGGAGAACT GCGGCTATCG CAGGTGCCGC GCGAGCGCG TACTGACAAC TAGCTGTGGT  
 CGGGAGAATG GTTAAGTTCC CCCCTCTTGA CGCCGATAGC GTCCACGGCG CGCTCGCCGC ATGACTGTTG ATCGACACCA

---

+2 N T L T C Y I K A R A A C R A A G L Q D C T M L V C G  
 6401 AACACCTCA CTGCTACAT CAAGGCCCGG GCAGCCTGTC GAGCCGAGG GCTCCAGGAC TGCAACATGC TCGTGTGG  
 TTGTGGGAGT GAACGATGTA GTTCCGGGCC CGTCGGACAG CTCGGCTCC CGAGGTCTG ACGTGGTACG AGCACACACC

---

+2 D D L V V I C E S A G V Q E D A A S L R A F T E A M  
 6481 CGACGACTTA GTCGTTATCT GTGAAAGCGC GGGGGTCCAG GAGGACGCGG CGAGCCTGAG AGCCTTCACG GAGGCTATGA  
 GCTGCTGAAT CAGCAATAGA CACTTCGCG CCCCCAGGTC CTCCCTGCCTC GCTCGGACTC TCGGAAGTGC CTCCGATACT

---

+2 T R Y S A P P G D P P Q P E Y D L E L I T S C S S N V  
 6561 CCAGGTACTC CGCCCCCCCCCT GGGGACCCCG CACAACCGA ATACGACTTG GAGCTATAA CATCATGTC CTCCAAACGTC  
 GGTCCATGAG GCGGGGGGGGA CCCCTGGGG GTGTTGGTCT TATGCTAAC CTCGAGTATT GTAGTACGAG GAGGTTGCAC

---

+2 S V A H D G A G K R V Y Y L T R D P T T P L A R A A W  
 6641 TCAGTCGCC ACGACGGCGC TGGAAAGAGG GTCTACTACC TCACCCGTGA CCCTACAACC CCCCTCGCGA GAGCTGCGTG  
 AGTCAGCGGG TGCTGCCCGC ACCTTTCTCC CAGATGATGG AGTGGGCAG GGGATGTTGG GGGGAGCGCT CTCGACGCAC

---

+2 E T A R H T P V N S W L G N I I M F A P T L W A R M  
 6721 GGAGACAGCA AGACACACTC CAGTCATTCTC CTGGCTAGGC AACATAATCA TGTTGCCCC CACACTGTGG GCGAGGATGA  
 CCTCTGTCGT TCTGTGTAG GACCGATCCG TTGTATTAGT ACAAACGGGG GTGTGACACC CGCTCCTACT

---

+2 I L M T H F F S V L I A R D Q L E Q A L D C E I Y G A  
 6801 TACTGATGAC CCATTCTTT AGCGTCTTTA TAGCCAGGGA CCAGCTTGAA CAGGCCCTCG ATTGCAGAT CTACGGGGCC  
 ATGACTACTG GGTAAAGAAA TCGCAGGAAT ATCGGTCCTT GGTCGAACCT GTCCGGGAGC TAACGCTCTA GATGCCCGCC

---

+2 C Y S I E P L D L P P I I Q R L H G L S A F S L H S Y  
 6881 TGCTACTCCA TAGAACCAACT GGATCTACCT CCAATCATTC AAAGACTCCA TGGCCTCAGC GCATTTCAC TCCACAGTTA  
 ACGATGAGGT ATCTGGTGA CCTAGATGGA GTTGTAGTAAG TTTCTGAGGT ACCGGAGTCG CGTAAAAGTG AGGTGTCAAT

---

+2 S P G E I N R V A A C L R K L G V P P L R A W R H R  
 6961 CTCTCCAGGT GAAATCAATA GGGTGGCGC ATGCCTCAGA AAACCTGGGG TACCGCCCTT GCGAGCTGG AGACACCGGG  
 GAGAGGTCCA CTTTAGTTAT CCCACCGCG TACGGAGTCT TTTGAACCCC ATGGCGGGAA CGCTCGAAC TCTGTGGCC

---

+2 A R S V R A R L L A R G G R A A I C G K Y L F N W A V  
 7041 CCCGGAGCGT CCGCGCTAGG CTTCTGGCCA GAGGAGGCAG GGCTGCCATA TGTGGCAAGT ACCTCTCAA CTGGCGAGTA  
 GGGCCTCGCA GGCGCGATCC GAAGACCGGT CTCCCTCCGTC CCGACGGTAT ACACCGTTCA TGGAGAAGTT GACCCGTCAT

## FIGURE 5 - Page 7

+2 R T K L K L T P I A A A G Q L D L S G W F T A G Y S G  
 7121 AGAACAAAGC TCAAACAC TCCAATAGCG GCCGCTGGCC AGCTGGACTT GTCCGGCTGG TTCACGGCTG GCTACAGCGG  
 TCTTGTTCG AGTTGAGTG AGTTATCGC CGGGACCGG TCGACCTGAA CAGGCCGACC AAGTGCCGAC CGATGTCGCC

+2 G D I Y H S V S H A R P R W I W F C L L L L A A G V  
 7201 GGGAGACATT TATCACAGCG TGTCTCATGC CGGGCCCCGC TGGATCTGGT TTTGCCTACT CCTGCTTGCT GCAGGGTAG  
 CCCCTGTAA ATAGTGTGCG ACAGAGTACG GGCGGGGGCG ACCTAGACCA AAACGGATGA GGACGAACGA CGTCCCCATC

+2 G I Y L L P N R  
 7281 GCATCTACCT CCTCCCCAAC CGATGAAGGT TGGGGTAAAC ACTCCGGCCT AAAAAAAA AAAAATCTAG AAAGGCCGCG  
 CGTAGATGGA GGAGGGGTTG GCTACTTCCA ACCCCATTG TGAGGCCGGA TTTTTTTTT TTTTAGATC TTTCCGCGCG

BamHI      MluI

~~~~~      ~~~~~

7361 CAAGATATCA AGGATCCACT ACGCGTTAGA GCTCGCTGAT CAGCCTCGAC TGTGCCCTCT AGTTGCCAGC CATCTGTTGT  
 GTTCTATAGT TCCTAGGTGA TGCGCAATCT CGAGCGACTA GTCGGAGCTG ACACGGAAGA TCAACGGTCG GTAGACAACA

7441 TTGCCCTCC CCCGTGCCTT CCTTGACCTT GGAAGGTGCC ACTCCCACGT TCCTTCCTA ATAAAATGAG GAAATTGCAT  
 AACGGGAGG GGGCACGGAA GGAACGGGA CCTTCCACGG TGAGGGTAC AGGAAAGGAT TATTTACTC CTTAACGTA

7521 CGCATTGTCT GAGTAGGTGT CATTCTATTG TGGGGGGTGG GGTGGGGCAG GACAGCAAGG GGGAGGATTG GGAAGACAAT  
 GCGTAACAGA CTCATCCACA GTAAGATAAG ACCCCCCACC CCACCCCGTC CTGTCGTTCC CCCTCCTAAC CCTTCTGTAA

7601 AGCAGGCATG CTGGGGAGCT CTTCCGCTTC CTCCGCTACT GACTCGCTGC GCTCGGTCGT TCGGCTGCCG CGAGCGGTAT  
 TCGTCGTTAC GACCCCTCGA GAAGGCGAAG GAGCGAGTGA CTGAGCGACG CGAGCCAGCA AGCCGACGCC GCTCGCCATA

7681 CAGCTCACTC AAAGGCCGTA ATACGGTTAT CCACAGAAC AGGGGATAAC GCAGGAAAGA ACATGTGAGC AAAAGGCCAG  
 GTCGAGTGAG TTTCCGCAAT TATGCCATA GGTGCTTAG TCCCCTATG CGTCCTTCT TGTACACTCG TTTTCCGTC

7761 CAAAAGGCCA GGAACCGTAA AAAGGCCGCG TTGCTGGCGT TTTCCATAG GCTCCGCCCG CCTGACGAGC ATCACAAAAA  
 GTTTCCGGT CCTTGGCATT TTTCCGGCGC AACGACCGCA AAAAGGTATC CGAGGCCGGG GGACTGCTCG TAGTGTTTTT

7841 TCGACGCTCA AGTCAGAGGT GGCAGAACCC GACAGGACTA TAAAGATACC AGGCCTTCC CCCTGGAAGC TCCCTCGTGC  
 AGCTCGAGT TCAGTCTCCA CCGCTTGGG CTGTCCTGAT ATTTCTATGG TCCGCAAAGG GGGACCTTCG AGGGAGCAG

7921 GCTCTCCTGT TCCGACCTG CCGCTTACCG GATACTGTC CGCCTTCTC CCTCAGGAA GCGTGGCGCT TTCTCAATGC  
 CGAGAGGACA AGGCTGGAC GCGAATGGC CTATGGACAG GCGGAAAGAG GGAAGCCCTT CGCACCGCGA AAGAGTTACG

8001 TCACGCTGTA GGTATCTCAG TTGGTGTAG GTGTCGCTG CCAAGCTGGG CTGTCGAC GAACCCCCCG TTCAGCCGAA  
 AGTGCAGACAT CCATAGAGTC AAGCCACATC CAGCAAGCGA GGTCGACCC GACACACGTG CTTGGGGGGC AAGTGGGCT

8081 CCGCTGCAGCC TTATCCGTA ACTATCGTCT TGAGTCAAAC CCGGTAAGAC ACGACTTATC GCCACTGGCA GCAGCCACTG  
 GCGACGCCGG AATAGGCCAT TGATAGCAGA ACTCAGGTG GGCCATTCTG TGCTGAATAG CGGTGACCGT CGTCGGTAC

8161 GTAACAGGAT TAGCAGAGCG AGGTATGTAG GCGGTGCTAC AGAGTTCTG AAGTGGTGGC CTAACACTGG CTACACTAGA  
 CATTGTCCTA ATCGTCTCGC TCCATACATC CGCCACGATG TCTCAAGAAC TTCACCCACCG GATTGATGCC GATGTGATCT

8241 AGGACAGTAT TTGGTATCTG CGCTCTGCTG AAGCCAGTTA CCTCAGGAA AAGAGTTGGT AGCTCTTGAT CGGGCAAACA  
 TCCTGTCATA AACCATAGAC GCGAGACGAC TTGGTCAAT GGAAGCCTT TTCTCAACCA TCGAGAACTA GGCGTTTGT

8321 AACCAACCGCT GGTAGCGGTG GTTTTTGT TTGCAAGCAG CAGATTACGC GCAGAAAAAAA AGGATCTCAA GAAGATCCTT  
 TTGGTGGCGA CCATCGCCAC CAAAAAAACA AACGTTGTCG GTCTAATGCG CGTCTTTTT TCCTAGAGTT CTTCTAGGAA

8401 TGATCTTTTC TACGGGGTCT GACGCTCAGT GGAAACGAAA CTCACGTTAA GGGATTTGG TCATGAGATT ATCAAAAGG  
 ACTAGAAAAG ATGCCAGA CTGCGAGTC CCTTGCTTT GAGTGCATT CCCTAAAACC AGTACTCTAA TAGTTTCC

## FIGURE 5 - Page 8

8481 ATCTTCACCT AGATCCTTTT AAATTAAAAA TGAAGTTTA AATCAATCTA AAGTATATAT GAGTAAACCTT GGTCTGACAG  
TAGAAGTGG A TCTAGGAAAA TTTAATTTC ACTTCAAAAT TTAGTTAGAT TTCAATATATA CTCATTGAA CCAGACTGTC

8561 TTACCAATGC TTAATCAGTG AGGCACCTAT CTCAGCGATC TGTCTATTTC GTTCATCCAT AGTTGCCTGA CTCCCCGTCG  
AATGGTTACG AATTAGTCAC TCCGTGGATA GAGTCGCTAG ACAGATAAAG CAAGTAGGTA TCAACGGACT GAGGGCAGC

8641 TGTAGATAAC TACGATACGG GAGGGCTTAC CATCTGGCCC CAGTGTGCA ATGATACCGC GAGACCCACG CTCACCGGCT  
ACATCTATTG ATGCTATGCC CTCCCGAATG GTAGACCGGG GTCACGACGT TACTATGGCG CTCTGGGTGC GAGTGGCCGA

8721 CCAGATTAT CAGCAATAAA CCAGCCAGCC GGAAGGGCCG AGCGCAGAAG TGTTGCCTGCA ACTTTATCCG CCTCCATCCA  
GGTCTAAATA GTCGTTATTG GGTCGGTCGG CCTTCCCAGC TCGCGTCTC ACCAGGACGT TGAAATAGGC GGAGGTAGGT

8801 GTCTATTAAT TGTTGCCGGG AAGCTAGAGT AAGTAGTTCG CCAGTTAATA GTTTGCCTGAA CGTTGTTGCC ATTGCTACAG  
CAGATAATT ACAACGGCCC TTCGATCTCA TTCATCAAGC GGTCAATTAT CAAACCGT GCAACACGG TAACGATGTC

8881 GCATCGGGT GTCACGCTCG TCGTTGGTA TGGCTTCATT CAGCTCCGGT TCCCAACGAT CAAGGGCGAGT TACATGATCC  
CGTAGCACCA CAGTGCAGAC AGCAAACCAT ACCGAAGTAA GTCGAGGCCA AGGGTTGCTA GTTCCGCTCA ATGTACTAGG

8961 CCCATGTTGT GCAAAAAAGC GGTTAGCTCC TTGGTCCTC CGATCGTTGT CAGAAGTAAG TTGGCCCGAG TGTTATCACT  
GGGTACAACA CGTTTTTCG CCAATCGAGG AAGCCAGGAG GCTAGCAACA GTCTCATTC AACCGCGTC ACAATAGTGA

9041 CATGGTTATG GCAGCACTGC ATAATTCTCT TACTGTCTG CCATCCGAA GATGTTTTC TGTGACTGGT GAGTACTCAA  
GTACCAATAC CGTCGTGACG TATTAAGAGA ATGACAGTAC GGTAGGCATT CTACGAAAAG ACACTGACCA CTCATGAGTT

9121 CCAAGTCATT CTGAGAATAG TGTATGCCGC GACCGAGTTG CTCTGCCCG GCGTCAATAC GGGATAATAC CGCGCCACAT  
GGTCAGTAA GACTCTTATC ACATACGCCG CTGGCTCAAC GAGAACGGC CGCAGTTATG CCCTATTATG GCGCGGTGTA

9201 AGCAGAACTT TAAAAGTGCT CATCATTGGA AAACGTTCTT CGGGGCGAAA ACTCTCAAGG ATCTTACCGC TGTTGAGATC  
TCGTCTGAA ATTTCACGA GTAGTAACCT TTGCAAGAA GCCCCGCTT TGAGAGTCC TAGAATGGCG ACAACTCTAG

9281 CAGTTCGATG TAACCCACTC GTGCACCCAA CTGATCTTC GAATCTTTA CTTTACCCAG CGTTTCTGGG TGAGCAAAAAA  
GTCAAGCTAC ATTGGGTGAG CACGTGGTT GACTAGAAGT CGTAGAAAAT GAAAGTGGTC GCAAAGACCC ACTCGTTTT

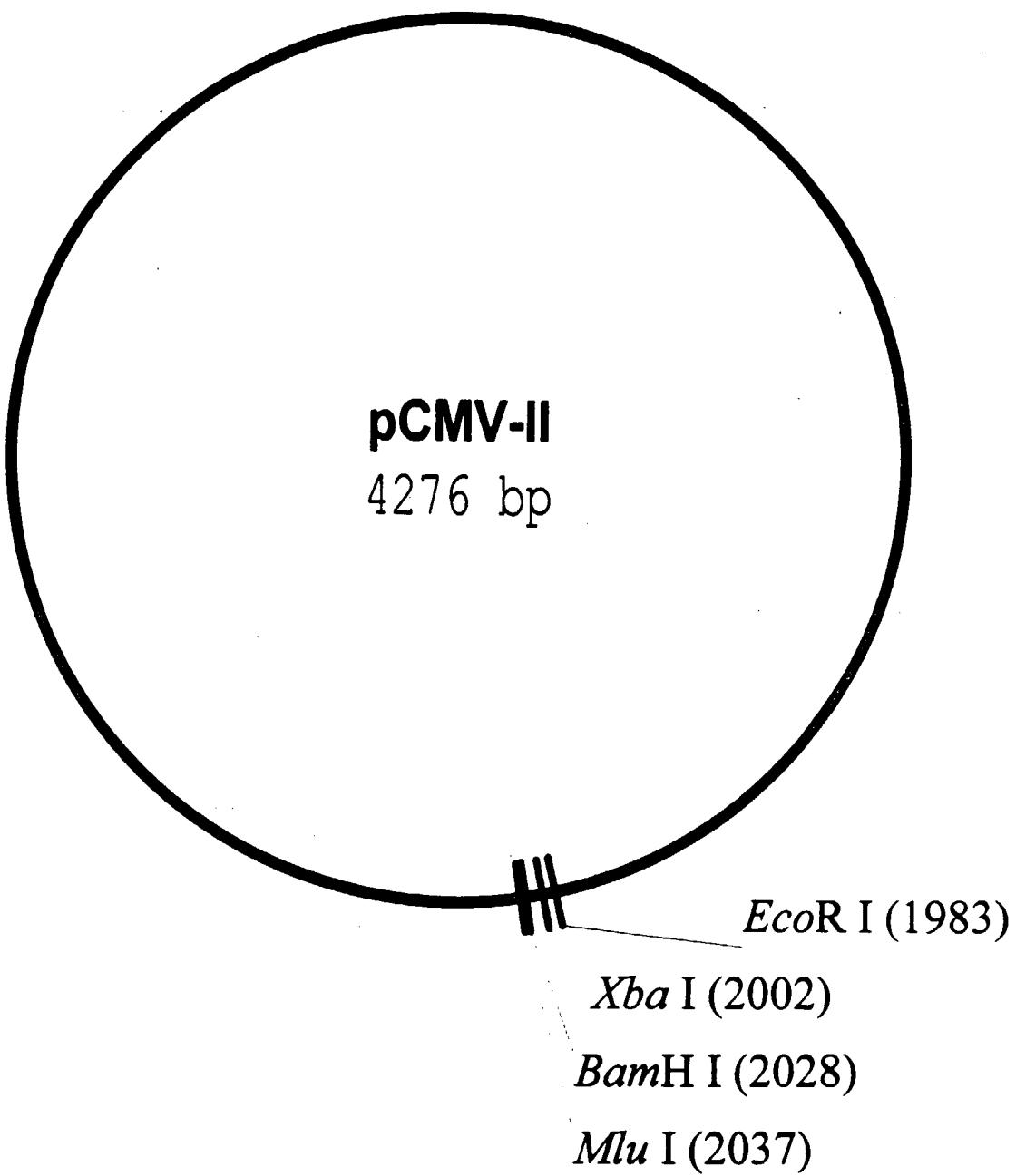
9361 CAGGAAGGCA AAATGCCGCA AAAAGGGAA TAAGGGCGAC ACGGAAATGT TGAATACTCA TACTCTTCCT TTTCAATAT  
GTCCTTCGT TTTACGGCGT TTTTCCCTT ATTCCCGCTG TGCCTTACA ACTTATGAGT ATGAGAAGGA AAAAGTTATA

9441 TATTGAAGCA TTTATCAGGG TTATTGTCTC ATGAGCGGAT ACATATTGA ATGTATTTAG AAAATAAAC AAATAGGGGT  
ATAACTTCGT AAATAGTCCC AATAACAGAG TACTCGCCTA TGTATAACT TACATAAATC TTTTATTTG TTTATCCCCA

9521 TCCGCGCACA TTTCGGAA AAGTGCCACC TGACGTCTAA GAAACCATTA TTATCATGAC ATTAACCTAT AAAATAAGGC  
AGGCACGTGTT AAAGGGGCTT TTCACGGTGG ACTGCAGATT CTTGGTAAT AATAGTACTG TAATTGGATA TTTTATCCG

9601 GTATCACGAG GCCCTTTCGT C  
CATAGTGTCT CGGGAAAGCA G

FIGURE 6



## FIGURE 7 - Page 1

1 TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG GAGACGGTCA CAGCTTGTCT GTAAGCGGAT  
AGCGCGAAA GCCACTACTG CCACTTTGG AGACTGTGTA CGTCGAGGGC CTCTGCCAGT GTCGAACAGA CATTGCCATA

---

81 GCCGGGAGCA GACAAGCCCG TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG CGGCATCAGA  
CGGCCCTCGT CTGTTCGGGC AGTCCCGCGC AGTCGCCAC AACCGCCAC AGCCCGACC GAATTGATAC GCCGTAGTCT

---

161 GCAGATTGTA CTGAGAGTGC ACCATATGAA GCTTTTGCA AAAGCCTAGG CCTCCAAAAA AGCCTCCTCA CTACTTCTGG  
CGTCTAACAT GACTCTCACG TGGTATACTT CGAAAAACGT TTTGGATCC GGAGGTTTT TCGGAGGAGT GATGAAGACC

---

241 AATAGCTCAG AGGCCGAGGC GGCCTCGGCC TCTGCATAAA TAAAAAAAAT TAGTCAGCCA TGGGGCGGAG AATGGGCAGA  
TTATCGAGTC TCCGGCTCCG CGCGAGCCGG AGACGTATTT ATTTTTTTA ATCAGTCGGT ACCCCGCCTC TTACCCGCCT

---

321 ACTGGGCGGG GAGGGAATTA TTGGCTATTG GCCATTGCAT ACgttGTATC TATATCATAA TATGTACATT TATATTGGCT  
TGACCCGCC CGCCCTTAAT AACCGATAAC CGGTAAACGTA TGCAACATAG ATATAGTATT ATACATGTAA ATATAACCGA

---

401 CATGTCCAAT ATGACCGCCA TGTTGACATT GATATTGAC TAGTTATTAA TAGTAATCAA TTACGGGTC ATTAGTTCA  
GTACAGGTTA TACTGGCGGT ACAACTGTAA CTAATAACTG ATCAATAATT ATCATTAGTT AATGCCAG TAATCAAGTA

---

481 AGCCCATATA TGAGGTTCCG CGTTACATAA CTTACGGTAA ATGGCCGCC TGGCTGACCG CCCAACGACC CCCGCCATT  
TCGGGTATAT ACCTCAAGGC GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC GGGTTGCTGG GGGCGGGTAA

---

561 GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCATA GGGACTTCC ATTGACGTCA ATGGGTGGAG TATTTACGGT  
CTGCAGTTAT TACTGCATAC AAGGGTATCA TTGGGTTAT CCCTGAAAGG TAACTGCAGT TACCCACCTC ATAAATGCCA

---

641 AAACTGCCCA CTGGCAGTA CATCAAGTGT ATCATATGCC AAGTCCGCC CCTATTGACG TCAATGACGG TAAATGGCCC  
TTGACGGGT GAACCGTCAT GTAGTTACA TAGTATACGG TTCAGGGGG GGATAACTGC AGTTACTGCC ATTTACCGGG

---

721 GCCTGGCATT ATGCCAGTA CATGACCTTA CGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC  
CGGACCGTAA TACGGGTAT GTACTGGAAT GCCCTGAAAG GATGAACCGT CATGTAGATG CATAATCAGT AGCGATAATG

---

801 CATGGTGATG CGGTTTGGC AGTACACCAA TGGCGTGGG TAGCGGTTTG ACTCACGGGG ATTTCCAAGT CTCCACCCCA  
GTACCACTAC GCCAAAACCG TCATGTGGTT ACCCGCACCT ATGCCAACAC TGAGTGCCTC TAAAGGTTCA GAGGTGGGGT

---

881 TTGACGTCAA TGGAGGTTG TTTGGCACC AAAATCAACG GGACTTCCA AAATGTGTA ATAACCCCGC CCCGTTGACG  
AACTGCAGTT ACCCTCAAAC AAAACCGTGG TTTAGTTGC CCTGAAAGGT TTACAGCAT TATTGGGGCG GGGCAACTGC

---

961 CAAATGGCG GTAGGGGTGT ACGGTGGAG GTCTATATAA GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG  
GTTTACCCGC CATCCGCACA TGCCACCCCTC CAGATATATT CGTCTCGAGC AAATCACTTG GCAGTCTAGC GGACCTCTGC

---

1041 CCATCCACGC TGTTTGACC TCCATAGAAAG ACACCGGGAC CGATCCAGCC TCCGCGGCCG GGAACGGTGC ATTGGAACGC  
GGTAGGTGCG ACAAAACCTGG AGGTATCTTC TGTGGCCCTG GCTAGGTGG AGGCGCCGGC CCTTGCCACG TAACCTTGCG

---

1121 GGATTCCCCG TGCCAAGAGT GACGTAAGTA CCCCTATAG ACTCTATAGG CACACCCCTT TGGCTCTTAT GCATGCTATA  
CCTAAGGGGC ACGGTTCTCA CTGCATTCTCAT GGCGGATATC TGAGATATCC GTGTGGGAA ACCGAGAATA CGTACGATAT

---

1201 CTGTTTTGG CTTGGGGCCT ATACACCCCC GCTCCTTATG CTATAGGTGA TGGTATAGCT TAGCCTATAG GTGTGGGTTA  
GACAAAACCG GAACCCCGGA TATGTGGGG CGAGGAATAC GATATCCACT ACCATATCGA ATCGGATATC CACACCAAT

---

1281 TTGACCATTA TTGACCACTC CCCTATTGGT GACGATACTT TCCATTACTA ATCCATAACA TGGCTCTTGC CCACAACTAT  
AACTGTAAT AACTGGTGG AGGATAACCA CTGCTATGAA AGGTAATGAT TAGGTATTGT ACCGAGAAC GGTGTTGATA

---

1361 CTCTATTGGC TATATGCCAA TACTCTGTCC TTCAGAGACT GACACGGACT CTGTATTTT ACAGGATGGG GTCCATTAT  
GAGATAACCG ATATACGGTT ATGAGACAGG AAGTCTCTGA CTGTGCCCTGA GACATAAAA TGTCCTACCC CAGGTAAATA

---

1441 TATTTACAAA TTCACATATA CAACAAACGCC GTCCCCCGTG CCCGCAGTTT TTATTAACAA TAGCGTGGGA TCTCCGACAT  
ATAATGTTT AAGTGTATAT GTTGTGCGG CAGGGGGCAC GGGCGTCAA AATAATTGT ATCGCACCC AGAGGCTGTA

## FIGURE 7 - Page 2

1521 CTCGGGTACG TGTTCGGAC ATGGGCTCTT CTCCGGTAGC GGCGGAGCTT CCACATCCGA GCCCTGGTCC CATCCGTCCA  
GAGCCCATGC ACAAGGCCTG TACCCGAGAA GAGGCCATCG CGGCCCTCGAA GGTGTAGGCT CGGGACCAGG GTAGGCAGGT

---

1601 GCGGCTCATG GTGCGCTCGC AGCTCCTGTC TCCTAACAGT GGAGGCCAGA CTTAGGCACA GCACAATGCC CACCACCAC  
CGCCGAGTAC CAGCGAGCCG TCGAGGAACG AGGATTGTCA CCTCCGGTCT GAATCCGTGT CGTGTACGG GTGGTGGTGG

---

1681 AGTGTGCCGC ACAAGGCCGT GGCAGTAGGG TATGTGTCTG AAAATGAGCT CGGAGATTGG GCTCGCACCT GGACGCAGAT  
TCACACGGCG TGTTCGGCA CCGCCATCCC ATACACAGAC TTTTACTCGA GCCTCTAACCG CGAGCGTGGA CCTGCGTCTA

---

1761 GGAAGACTTA AGGCAGCGGC AGAAGAAGAT GCAGGCAGCT GAGTTGTTGT ATTCTGATAA GAGTCAGAGG TAACTCCCGT  
CCTTCTGAAT TCCGTCGCCG TCTTCTTCTA CGTCCGTCGA CTCAACAACA TAAGACTATT CTCAGTCTCC ATTGAGGGCA

---

1841 TGCAGGTGCTG TTAACGGTGG AGGGCAGTGT AGTCTGAGCA GTACTCGTTG CTGCCCGCG CGCCACCAGA CATAATAGCT  
ACGCCACGAC AATTGCCACC TCCCCTCACAC TCAAGACTCGT CATGAGCAAC GACGGCGCGC GCGGTGGTCT GTATTATCGA

EcoRI

~~~~~

1921 GACAGACTAA CAGACTGTTC CTTCCATGG GTCTTTCTG CAGTCACCGT CGTCGACCTA AGAATTAGA CTCGAGCAAG  
CTGTCTGATT GTCTGACAAG GAAAGGTACC CAGAAAAGAC GTCAGTGGCA GCAGCTGGAT TCTTAAGTCT GAGCTCGTTC

XbaI

BamHI

MluI

~~~~~

~~~~~

~~~~~

2001 TCTAGAAAGG CGCGCCAAGA TATCAAGGAT CCAACTACCGT TTAGAGCTCG CTGATCAGCC TCGACTGTGC CTTCTAGTTG  
AGATCTTCC CGCGGGTTCT ATAGTTCTA GGTGATGCGC AATCTCGAGC GACTAGTCGG AGCTGACACG GAAGATCAAC

2081 CCAGCCATCT GTGTTTGCC CCTCCCCGT GCCTTCCTTG ACCCTGGAAG GTGCCACTCC CACTGTCCTT TCCTAATAAA  
GGTCGGTAGA CAACAAACGG GGAGGGGGCA CGGAAGGAAC TGGGACCTTC CACGGTGAGG GTGACAGGAA AGGATTATTT

2161 ATGAGGAAAT TGCATCGCAT TGTCTGAGTA GGTGTCATTG TATTCTGGGG GGTGGGGTGG GGCAGGACAG CAAGGGGGAG  
TACTCCTTA ACGTAGCGTA ACAGACTCAT CCACAGTAAG ATAAGACCCC CCACCCACC CCGTCCTGTC GTTCCCCCTC

2241 GATTGGGAAG ACAATAGCAG GCATGCTGGG GAGCTCTTC GCTTCCTCGC TCACTGACTC GCTGCGCTCG GTCGTCGGC  
CTAACCTTC TGTTATCGTC CGTACGACCC CTCGAGAAGG CGAAGGAGCG AGTGAATGAG CGACCGGAGC CAGCAAGCCG

2321 TGCAGCGAGC GGTATCAGCT CACTCAAAGG CGGTAAATACG GTTATCCACA GAATCAGGGG ATAACGCAGG AAAGAACATG  
ACGCCGCTCG CCATAGTCGA GTGAGTTCC GCCATTATGC CAATAGGTGT CTTAGTCCCC TATTGCGTCC TTTCTGTAC

2401 TGAGCAAAAG GCCAGCAAA GGCCAGGAAC CGTAAAAAGG CCGCGTTGCT GGCGTTTTTC CATAGGCTCC GCCCCCTGA  
ACTCGTTTC CGGTCGTTT CCGGTCTTG GCATTTTCC GGCGCAACGA CGCAGGAAAG GTATCCGAGG CGGGGGGACT

2481 CGAGCATCAC AAAATCGAC GCTCAAGTCA GAGGTGGCGA AACCCGACAG GACTATAAG ATACCAGGCG TTTCCCCCTG  
GCTCGTAGTG TTTTAGCTG CGAGTTCAGT CTCCACCGCT TTGGGCTGTC CTGATATTTC TATGGTCCGC AAAGGGGGAC

2561 GAAGCTCCCT CGTGCCTCT CCTGTTCCGA CCCTGCCGCT TACCGGATAC CTGCCGCTT TTCTCCCTTC GGGAGCGTG  
CTTCGAGGGGA GCACCGGAGA GGACAAGGCT GGGACGGCGA ATGGCCTATG GACAGCGGA AAGAGGGAAAG CCCTCGCAC

2641 GCGCTTCTC AATGCTCAGC CTGTTAGGTAT CTCAGTCGG TGTTAGTCGT TCGCTCCAAG CTGGGCTGTC TGCACGAACC  
CGCGAAAGAG TTACGAGTGC GACATCCATA GAGTCAAGCC ACATCCAGCA AGCGAGGTTC GACCCGACAC ACGTGCTTGG

2721 CCCCCTTAG CCCCACCGCT GCGCCTTATC CGGTAACTAT CGTCTTGAGT CCAACCCGGT AAGACACGAC TTATCGCCAC  
GGGGCAAGTC GGGCTGGCGA CGCGGAATAG GCCATTGATA GCAGAACTCA GGTTGGCCA TTCTGTGCTG AATAGCGGTG

2801 TGGCAGCAGC CACTGGTAAC AGGATTAGCA GAGCGAGGTA TGTTAGGCCTT GCTACAGAGT TCTTGAAGTG GTGGCCTAAC  
ACCGTCGTCG GTGACCATTG TCCTAACCGT CTCGCTCCAT ACATCCGCCA CGATGTCTCA AGAACTTCAC CACCGGATTG

2881 TACGGCTACA CTAGAAGGAC AGTATTGAGT ATCTGCGCTC TGCTGAAGCC AGTACCTTC GGAAAAAGAG TTGGTAGCTC  
ATGCCGATGT GATCTTCCCTG TCATAAACCA TAGACCGAG ACGACTTCGG TCAATGAAAG CCTTTTCTC AACCATCGAG

## FIGURE 7 - Page 3

2961 TTGATCCGGC AAACAAACCA CCGCTGGTAG CGGTGGTTT TTTGTTGCA AGCAGCAGAT TACGCGCAGA AAAAAAGGAT  
AACTAGGCCG TTTGTTGGT GGCGACCATC GCCACCAAA AAACAAACGT TCGTCGTCTA ATGCGCGTCT TTTTTCCCTA

3041 CTCAAGAAGA TCCCTTGATC TTTTCTACGG GGCTGTACGC TCAGTGGAAC GAAAACCTAC GTTAAGGGAT TTTGGTCATG  
GAGTTCTTCT AGGAAACTAG AAAAGATGCC CCAGACTGCG AGTCACCTG CTTTGAGTG CAATTCCCTA AAACCAGTAC

3121 AGATTATCAA AAAGGATCTT CACCTAGATC CTTTAAATT AAAATGAAG TTTAAATCA ATCTAAAGTA TATATGAGTA  
TCTAATAGTT TTTCTAGAA GTGGATCTAG GAAAATTAA TTTTACTTC AAAATTAGT TAGATTCAT ATATACTCAT

3201 AACTTGGTCT GACAGTTACC AATGCTTAAT CAGTGAGGCA CCTATCTAG CGATCTGTCT ATTTCGTTCA TCCATAGTTG  
TTGAACCGA CTGTCATGG TTACGAATTA GTCACCTCGT GGATAGAGTC GCTAGACAGA TAAAGCAAGT AGGTATCAAC

3281 CCTGACTCCC CGTCGTGTAG ATAACACGA TACGGGAGGG CTTACCATCT GGCCCCAGTG CTGCAATGAT ACCCGCAGAC  
GGACTGAGGG GCAGCACATC TATTGATGCT ATGCCCTCCC GAATGGTAGA CCGGGTCAC GACGTTACTA TGGCGCTCTG

3361 CCACGCTCAC CGGCTCCAGA TTTATCAGCA ATAAACCAGC CAGCCGGAAG GGCGCAGCGC AGAAGTGGTC CTGCAACTT  
GGTGCAGTG GCCGAGGTCT AAATAGTCGT TATTGGTCG GTCGGCCTTC CCGGCTCGCG TCTTCACCAAG GACGTTGAAA

3441 ATCCGCCTCC ATCCAGTCTA TTAATTGTTG CCGGGAAAGCT AGAGTAAGTA GTTCGCCAGT TAATAGTTG CGCAACGTTG  
TAGGCGGAGG TAGGTAGAT AATTAACAAAC GGCCCTTCGA TCTCATTCTAT CAAGCGGTCA ATTATCAAAC GCGTTGCAAC

3521 TTGCCATTGC TACAGGCATC GTGGTGTAC GCTCGTGTGTT TGGTATGGCT TCATTCAGCT CCGGTTCCCA ACGATCAAGG  
AACGGTAACG ATGTCCGTAG CACCACAGTG CGAGCAGCAA ACCATACCGA AGTAAGTCGA GGCAAGGGT TGCTAGTTCC

3601 CGAGTTACAT GATCCCCAT GTTGTGCAAA AAAGCGGTTA GCTCCTTCGG TCCTCCGATC GTTGTCAAGAA GTAAGTGGC  
GCTCAATGTA CTAGGGGTA CAACACGTTT TTTCGCAAT CGAGGAAGCC AGGAGGCTAG CAACAGTCTT CATTCAACCG

3681 CGCAGTGTAA TCACTCATGG TTATGGCAGC ACTGCATAAT TCTCTTACTG TCATGCCATC CGTAAGATGC TTTTCTGTGA  
GCGTCACAAT AGTGAGTACC AATACCGTCG TGACGTATTA AGAGAATGAC AGTACGGTAG GCATTCTACG AAAAGACACT

3761 CTGGTGAGTA CTCAACCAAG TCATTCTGAG AATAGTGTAT GCGGCGACCG AGTTGCTCTT GCCCCGGCGTC AATACGGGAT  
GACCACTCAT GAGTTGGTTC AGTAAGACTC TTATCACATA CGCCGCTGGC TCAACGAGAA CGGGCGCGAG TTATGCCCTA

3841 AATACCGCGC CACATAGCAG AACTTTAAA GTGCTCATCA TTGGAAAACG TTCTCGGGG CGAAAACCTCT CAAGGATCTT  
TTATGGCGCG GTGTATGTC TTGAAATTTC CACGAGTAGT AACCTTTGC AAGAAGCCCC GCTTTGAGA GTTCCTAGAA

3921 ACCGCTGTTG AGATCCAGTT CGATGTAACC CACTCGTCA CCCAACTGAT CTTCAGCATC TTTTACTTTC ACCAGCGTT  
TGGCGACAAC TCTAGGTCAA GCTACATTGG GTGAGCACGT GGTTGACTA GAAGTCGTAG AAAATGAAAG TGGTCGCAAA

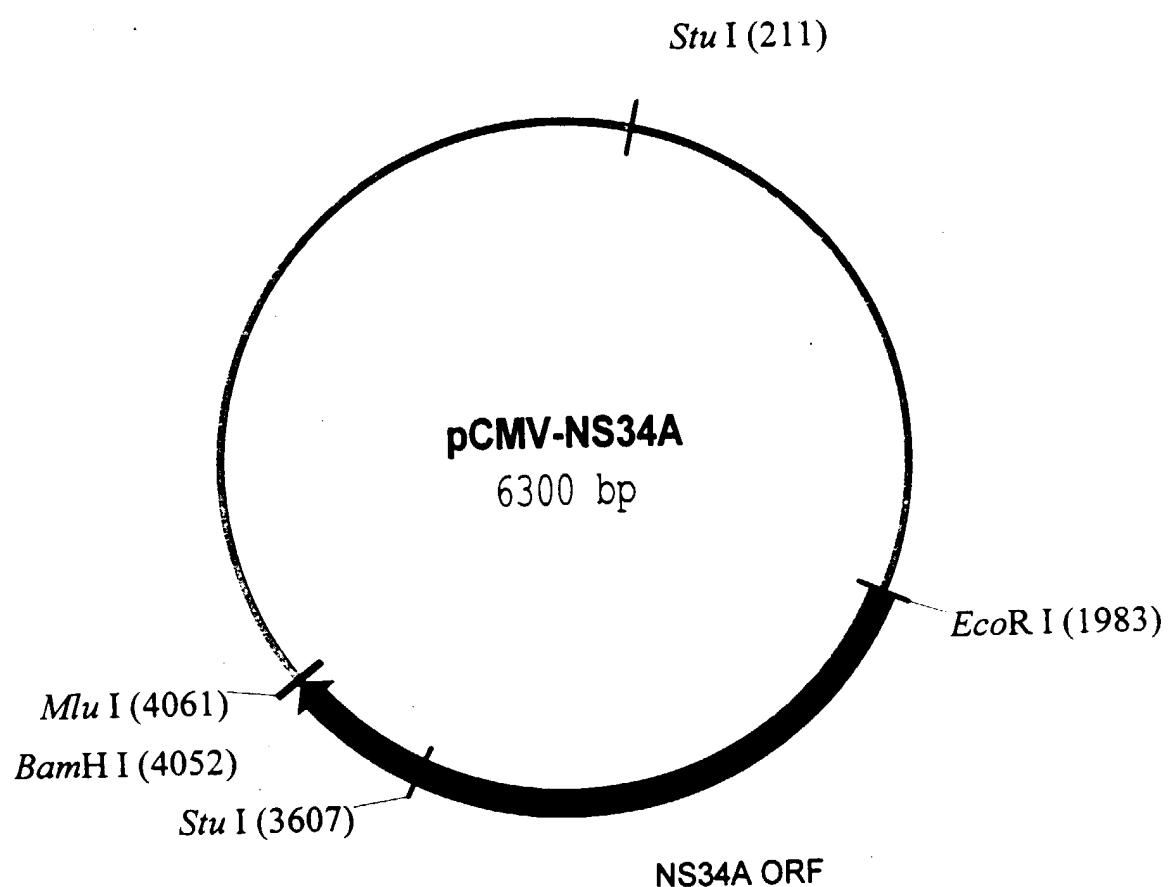
4001 CTGGGTGAGC AAAAAACAGGA AGGCAAAATG CCGCAAAAAA GGGATAAGG GCGACACCGA AATGTTGAAT ACTCATAC  
GACCCACTCG TTTTGTCTT TCCGTTTAC GGCGTTTTT CCCTTATTCC CGCTGTGCCT TTACAACCTTA TGAGTATGAG

4081 TTCCCTTTTC AATATTATTG AAGCATTAT CAGGGTTATT GTCTCATGAG CGGATACATA TTTGAATGTA TTTAGAAAAA  
AAGGAAAAAG TTATAATAAC TTCGTAATAA GTCCAATAA CAGAGTACTC GCCTATGTAT AAACCTTACAT AAATCTTTT

4161 TAAACAAATA GGGGTTCCGC GCACATTCC CCGAAAAGTG CCACCTGACG TCTAAGAAC CATTATTATC ATGACATTAA  
ATTGTTTAT CCCAAGGCG CGTGTAAAGG GGCTTTAC GGTGGACTGC AGATTCTTG GTAATAATAG TACTGTAATT

4241 CCTATAAAA TAGGCGTATC ACGAGGCCCT TTGTC  
GGATATTTC ATCCGCATAG TGCTCCGGGA AAGCAG

**FIGURE 8**



## FIGURE 9 - Page 1

1 TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG  
AGCGCGAAA GCCACTACTG CCACCTTTGG AGACTGTGTA CGTCGAGGGC

51 GAGACGGTCA CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCG  
CTCTGCCAGT GTCGAACAGA CATTGCCTA CGGCCCTCGT CTGTCGGGC

101 TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG  
AGTCCCGCGC AGTCGCCAC AACCGCCCAC AGCCCCGACC GAATTGATAC

151 CGGCATCAGA GCAGATTGTA CTGAGAGTGC ACCATATGAA GCTTTTGCA  
GCCGTAGTCT CGTCTAACAT GACTCTCACG TGGTATACTT CGAAAAACGT

## StuI

201 AAAGCCTAGG CCTCCAAAAA AGCCTCCTCA CTACTTCTGG AATAGCTCAG  
TTTCGGATCC GGAGGTTTT TCAGGAGGAGT GATGAAGACC TTATCGAGTC

251 AGGCCGAGGC GGCCTCGGCC TCTGCATAAA TAAAAAAAAT TAGTCAGCCA  
TCCGGCTCCG CGGGAGCCGG AGACGTATTT ATTTTTTTA ATCAGTCGGT

301 TGGGGCGGAG AATGGGCGGA ACTGGGCGGG GAGGGAATTA TTGGCTATTG  
ACCCCGCCTC TTACCCGCCCT TGACCCGCCCT CTCCTTAAT AACCGATAAC

351 GCCATTGCAT ACGTTGTATC TATATCATAA TATGTACATT TATATTGGCT  
CGGTAACGTA TGCAACATAG ATATAGTATT ATACATGTAA ATATAACCGA

401 CATGTCCAAT ATGACGCCA TGTGACATT GATTATTGAC TAGTTATTAA  
GTACAGGTTA TACTGGCGGT ACAACTGTAA CTAATAACTG ATCAATAATT

451 TAGTAATCAA TTACGGGTC ATTAGTCAT AGCCCATATA TGGAGTTCCG  
ATCATTAGTT AATGCCCGAG TAATCAAGTA TCAGGTATAT ACCTCAAGGC

501 CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC  
GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC GGGTTGCTGG

551 CCCGCCATT GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA  
GGCGGGTAA CTGCAGTTAT TACTGCATAC AAGGGTATCA TTGCGGTTAT

601 GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTTACGGT AAACTGCCA  
CCCTGAAAGG TAACTGCAGT TACCCACCTC ATAAATGCCA TTGACGGGT

651 CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTCCGCC CCTATTGACG  
GAACCGTCAT GTAGTTACA TAGTATACGG TTCAGGCAGG GGATAACTGC

701 TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCAGTA CATGACCTTA  
AGTTACTGCC ATTACCGGG CGGACCGTAA TACGGTCAT GTACTGGAAT

751 CGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC  
GCCCTGAAAG GATGAACCGT CATGTAGATG CATAATCAGT AGCGATAATG

801 CATGGTGATG CGGTTTGGC AGTACACCAA TGGGCGTGG TAGCGGTTTG  
GTACCACTAC GCCAAAACCG TCATGTGGTT ACCCGCACCT ATGCCAAAC

851 ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG  
TGAGTGGCCC TAAAGGTTCA GAGGTGGGGT AACTGCAGTT ACCCTCAAAC

## FIGURE 9 - Page 2

901 TTTGGCAC CAAATCAACG GGACTTTCCA AAATGTCGTA ATAACCCCGC  
 AAAACCGTGG TTTAGTTGC CCTGAAAGGT TTACAGCAT TATTGGGCG  
 951 CCCGTTGACG CAAATGGCG GTAGGCCTGT ACGGTGGGAG GTCTATATAA  
 GGGCAACTGC GTTACCCGC CATCCGCACA TGCCACCCCTC CAGATATATT  
 1001 GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG CCATCCACGC  
 CGTCTCGAGC AAATCACTTG GCAGTCTAGC GGACCTCTGC GGTAGGTGCG  
 1051 TGTGACCC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCCGGGGCCG  
 ACAAAACTGG AGGTATCTTC TGTGCCCTG GCTAGGTGG AGGGCCCGGC  
 1101 GGAACGGTGC ATTGGAACGC GGATTCCCCG TGCCAAGAGT GACGTAAGTA  
 CCTGCCACG TAACCTTGCG CCTAAGGGGC ACGGTTCTCA CTGCATTAT  
 1151 CCGCCTATAG ACTCTATAGG CACACCCCT TGGCTCTTAT GCATGCTATA  
 GCGGGATATC TGAGATATCC GTGTGGGAA ACCGAGAATA CGTACGATAT  
 1201 CTGTTTTGG CTTGGGGCCT ATACACCCCC GCTCCTTAT CTATAGGTGA  
 GACAAAAACC GAACCCCGGA TATGTGGGG CGAGGAATAC GATATCCACT  
 1251 TGGTATAGCT TAGCCTATAG GTGTGGGTTA TTGACCATTA TTGACCACTC  
 ACCATATCGA ATCGGATATC CACACCCAA AACTGGTAAT AACTGGTGAG  
 1301 CCCTATTGGT GACGATACTT TCCATTACTA ATCCATAACA TGGCTTTG  
 GGGATAACCA CTGCTATGAA AGGTAAATGAT TAGGTATTGT ACCGAGAAC  
 1351 CCACAACTAT CTCTATTGGC TATATGCCAA TACTCTGTCC TTCAGAGACT  
 GGTGTTGATA GAGATAACCG ATATACGTT ATGAGACAGG AAGTCTCTGA  
 1401 GACACGGACT CTGTATTTT ACAGGATGGG GTCCATTAT TATTACAAA  
 CTGTGCCTGA GACATAAAA TGTCCCTACCC CAGGTAATA ATAATGTTT  
 1451 TTCACATATA CAACAAACGCC GTCCCCCGTG CCCGCAGTT TTATTAACA  
 AAGTGTATAT GTTGTGCGG CAGGGGGCAC GGGCGTCAA AATAATTGT  
 1501 TAGCGTGGGA TCTCCGACAT CTCGGGTACG TGTTCCGGAC ATGGGCTCTT  
 ATCGCACCCCT AGAGGCTGTA GAGCCATGC ACAAGGCCTG TACCCGAGAA  
 1551 CTCCGGTAGC GGCAGGAGCTT CCACATCCGA GCCCTGGTCC CATCCGTCCA  
 GAGGCCATCG CCGCCTCGAA GGTGTAGGCT CGGGACCAAGG GTAGGCAGGT  
 1601 GCGGCTCATG GTCGCTCGGC AGCTCCTTCG TCCTAACAGT GGAGGCCAGA  
 CGCCGAGTAC CAGCAGCGG TCGAGGAACG AGGATTGTCA CCTCCGGTCT  
 1651 CTTAGGCACA GCACAATGCC CACCACCAAGTGTGCCGC ACAAGGCCGT  
 GAATCCGTGT CGTGTACGG GTGGTGGTGG TCACACGGCG TGTTCCGGCA  
 1701 GGCAGGTAGGG TATGTGTCTG AAAATGAGCT CGGAGATTGG GCTCGCACCT  
 CGGCCATCCC ATACACAGAC TTTACTCGA GCCTCTAACCG CAGCGTGGAA  
 1751 GGACGCAGAT GGAAGACTTA AGGCAGCGGC AGAAGAAGAT GCAGGCAGCT  
 CCTCGTCTA CCTTCTGAAT TCCGTCGCCG TCTTCTCTA CGTCCGTGCA  
 1801 GAGTTGTTGT ATTCTGATAA GAGTCAGAGG TAACTCCCGT TGCGGTGCTG  
 CTCAACAAACA TAAGACTATT CTCAGTCTCC ATTGAGGGCA AGGCCACGAC

PCMV-NS34A

FIGURE 9 - Page 3

1851 TTAACGGTGG AGGCAGTGT AGTCTGAGCA GTACTCGTTG CTGCCGCGCG  
AATTGCCACC TCCCGTCACA TCAGACTCGT CATGAGCAAC GACGGCGCGC

1901 CGCCACCAGA CATAATAGCT GACAGACTAA CAGACTGTTG CTTTCCATGG  
GCGGTGGTCT GTATTATCGA CTGTCTGATT GTCTGACAAG GAAAGGTACC

+2

M A P

EcoRI

~~~~~

1951 GTCTTTCTG CAGTCACCGT CGTCGACCTA AGAATTCAAC ATGGCGCCCA  
CAGAAAAGAC GTCACTGGCA GCAGCTGGAT TCTTAAGTGG TACCGCGGGT

+2 I T A Y A Q Q T R G L L G C I I T  
2001 TCACGGCGTA CGCCCAGCAG ACAAGGGGCC TCCTAGGGTG CATAATCAC  
AGTGGCGCAT CGGGTTCGTC TGTTCCCCGG AGGATCCCAC GTATTAGTGG

+2 S L T G R D K N Q V E G E V Q I V  
2051 AGCCTAAGTG GCCGGGACAA AAACCAAGTG GAGGGTGAGG TCCAGATTG  
TCGGATTGAC CGGGCCTGTT TTTGGTTCAC CTCCCACTCC AGGTCTAACCA

+2 S T A A Q T F L A T C I N G V C  
2101 GTCAACTGCT GCCAAACCT TCCTGGCAAC GTGCATCAAT GGGGTGTGCT  
CAGTTGACGA CGGGTTTGGGA AGGACCGTTG CACGTAGTTA CCCCACACGA

+2 W T V Y H G A G T R T I A S P K G  
2151 GGACTGTCTA CCACGGGGCC GGAACGAGGA CCATCGCGTC ACCCAAGGGT  
CCTGACAGAT GGTGCCCCGG CCTTGCTCCT GGTAGCGCAG TGGGTTCCCA

+2 P V I Q M Y T N V D Q D L V G W P  
2201 CCTGTCATCC AGATGTATAC CAATGTAGAC CAAGACCTTG TGGGCTGGCC  
GGACAGTAGG TCTACATATG GTTACATCTG GTTCTGGAAC ACCCGACCGG

+2 A S Q G T R S L T P C T C G S S  
2251 CGCTTCGCAA GGTACCCGCT CATTGACACC CTGCACTTGC GGCTCCTCGG  
GCGAAGCGTT CCATGGGCGA GTAACTGTGG GACGTGAACG CCGAGGAGCC

+2 D L Y L V T R H A D V I P V R R R  
2301 ACCTTTACCT GGTACAGGAGG CACGCCGATG TCATTCCCGT GCGCCGGCGG  
TGGAAATGGGA CCAGTGCTCC GTGCGGCTAC AGTAAGGGCA CGCGGGCCCGC

+2 G D S R G S L L S P R P I S Y L K  
2351 GGTGATAGCA GGGCAGCCT GCTGTCGCC CGGCCATT CCTACTTGAA  
CCACTATCGT CCCCCGTGGA CGACAGCGGG GCGGGTAAA GGATGAACCTT

+2 G S S G G P L L C P A G H A V G  
2401 AGGCTCCTCG GGGGGTCCGC TGTTGTGCC CGCGGGGCAC GCCGTGGCA  
TCCGAGGAGC CCCCCAGGCG ACAACACGGG GCGCCCGTG CGGCACCCGT

+2 I F R A A V C T R G V A K A V D F  
2451 TATTTAGGGC CGCGGTGTGC ACCCGTGGAG TGGCTAAGGC GGTGGACTTT  
ATAAAATCCCG GCGCCACACG TGGGCACCTC ACCGATTCCG CCACCTGAAA

+2 I P V E N L E T T M R S P V F T D  
2501 ATCCCTGTGG AGAACCTAGA GACAACCATG AGGTCCCCGG TGTTCACCGA  
TAGGGACACC TCTTGGATCT CTGTTGGTAC TCCAGGGGCC ACAAGTGCCT

pCMV-NS34A  
FIGURE 9 - Page 4

+2 N S S P P V V P Q S F Q V A H L  
 2551 TAACTCCTCT CCACCAAGTAG TGCCCCAGAG CTTCCAGGTG GCTCACCTCC  
 ATTGAGGAGA GGTGGTCATC ACGGGGTCTC GAAGGTCCAC CGAGTGGAGG

---

+2 H A P T G S G K S T K V P A A Y A  
 2601 ATGCTCCAC AGGCAGCGC AAAAGCACCA AGGTCCCAGC TGCAATATGCA  
 TACGAGGGTG TCCGTCGGCG TTTCTGTGGT TCCAGGGCCG ACGTATACTG

---

+2 A Q G Y K V L V L N P S V A A T L  
 2651 GCTCAGGGCT ATAAGGTGCT AGTACTCAAC CCCTCTGTG TGCAACACT  
 CGAGTCCCGA TATTCCACGA TCATGAGTTG GGGAGACAAC GACGTTGTGA

---

+2 G F G A Y M S K A H G I D P N I  
 2701 GGGCTTTGGT GCTTACATGT CCAAGGCTCA TGGGATCGAT CCTAACATCA  
 CCCGAAACCA CGAATGTACA GGTTCCGAGT ACCCTAGCTA GGATTGTAGT

---

+2 R T G V R T I T T G S P I T Y S T  
 2751 GGACCGGGGT GAGAACAAATT ACCACTGGCA GCCCCATCAC GTACTCCACC  
 CCTGGCCCCA CTCTTGTAA TGGTGACCGT CGGGTAGTG CATGAGGTGG

---

+2 Y G K F L A D G G C S G G A Y D I  
 2801 TACGGCAAGT TCCTTGCCGA CGGGGGGTGC TCGGGGGGGCG CTTATGACAT  
 ATGCCGTTCA AGGAACGGCT GCCGCCACG AGCCCCCCCAG GAATACTGTA

---

+2 I I C D E C H S T D A T S I L G  
 2851 AATAATTGT GACGAGTGCC ACTCCACGGA TGCCACATCC ATCTTGGGCA  
 TTATTAACCA CTGCTCACCG TGAGGTGCCT ACGGTGTAGG TAGAACCGT

---

+2 I G T V L D Q A E T A G A R L V V  
 2901 TTGGCACTGT CCTTGACCAA GCAGAGACTG CGGGGGCGAG ACTGGTTGTG  
 AACCGTGACA GGAACGGTT CGTCTCTGAC GCCCCCGCTC TGACCAACAC

---

+2 L A T A T P P G S V T V P H P N I  
 2951 CTCGCCACCG CCACCCCTCC GGGCTCCGTC ACTGTGCCCC ATCCAAACAT  
 GAGCGGTGGC GGTGGGGAGG CCCGAGGCAG TGACACGGGG TAGGGTTGTA

---

+2 E E V A L S T T G E I P F Y G K  
 3001 CGAGGAGGTT GCTCTGTCCA CCACCGGAGA GATCCCTTT TACGGCAAGG  
 GCTCCTCAA CGAGACAGGT GGTCGGCTCT CTAGGGAAAA ATGCCGTTCC

---

+2 A I P L E V I K G G R H L I F C H  
 3051 CTATCCCCCT CGAAGTAATC AAGGGGGGGAGA GACATCTCAT CTTCTGTCA  
 GATAGGGGGAGA GCTTCATTAG TTCCCCCCCCT CTGTAGAGTA GAAGACAGTA

---

+2 S K K K C D E L A A K L V A L G I  
 3101 TCAAAGAAGA AGTGCACGA ACTCGCCGCA AAGCTGGTCG CATTGGCAT  
 AGTTCTCTC TCACGCTGCT TGAGCGCGT TTCGACCAGC GTAACCCGTA

---

+2 N A V A Y Y R G L D V S V I P T  
 3151 CAATGCCGTG GCCTACTACC GCGGTCTTGA CGTGTCCGTC ATCCCGACCA  
 GTACGGCAC CGGATGATGG CGCCAGAAC GCACAGGCAG TAGGGCTGGT

---

+2 S G D V V V V A T D A L M T G Y T  
 3201 GCGCGATGT TGTCGTG TGCGACGATG CCCTCATGAC CGGCTATACC  
 CGCCGCTACA ACAGCAGCAC CGTGGCTAC GGGAGTACTG GCCGATATGG

pCMV-NS34A

FIGURE 9 - Page 5

+2 G D F D S V I D C N T C V T Q T V  
 3251 GGCCTTCG ACTCGGTGAT AGACTGCAAT ACGTGTGTCA CCCAGACAGT  
 CCGCTGAAGC TGAGCCACTA TCTGACGTTA TGCACACAGT GGGTCTGTCA

+2 D F S L D P T F T I E T I T L P  
 3301 CGATTCAGC CTTGACCTTA CCTTCACCAT TGAGACAATC ACGCTCCCC  
 GCTAAAGTCG GAACTGGGAT GGAAGTGGTA ACTCTGTTAG TGCAGGGGG

+2 Q D A V S R T Q R R G R T G R G K  
 3351 AAGATGCTGT CTCCCGCACT CAACGTCGGG GCAGGACTGG CAGGGGGAAG  
 TTCTACGACA GAGGGCGTGA GTTGCAGCCC CGTCCTGACC GTCCCCCTTC

+2 P G I Y R F V A P G E R P S G M F  
 3401 CCAGGCATCT ACAGATTGTG GCCACCGGGG GAGCGCCCT CGGGCATGTT  
 GGTCCGTAGA TGTCTAACACA CCGTGGCCCC CTCGCGGGGA GGCGTACAA

+2 D S S V L C E C Y D A G C A W Y  
 3451 CGACTCGTCC GTCTCTGTG AGTGTATGA CGCAGGCTGT GCTTGGTATG  
 GCTGAGCAGG CAGGAGACAC TCACGATACT GCGTCCGACA CGAACCATAC

+2 E L T P A E T T V R L R A Y M N T  
 3501 AGCTACGCC CGCCGAGACT ACAGTTAGGC TACGAGCGTA CATGAACACC  
 TCGAGTGCAGG GCGGCTCTGA TGTCAATCCG ATGCTCGCAT GTACTTGTGG

+2 P G L P V C Q D H L E F W E G V F  
 3551 CCGGGGCTTC CCGTGTGCCA GGACCATCTT GAATTTGGG AGGGCGTCTT  
 GGCCCCGAAG GGCACACGGT CCTGGTAGAA CTTAAACCC TCCCGCAGAA

+2 T G L T H I D A H F L S Q T K Q  
 StuI  
 ~~~~~~  
 3601 TACAGGCCTC ACTCATATAG ATGCCACTT TCTATCCCAG ACAAAGCAGA  
 ATGTCCGGAG TGAGTATATC TACGGGTGAA AGATAGGGTC TGTTCTGTCT

+2 S G E N L P Y L V A Y Q A T V C A  
 3651 GTGGGGAGAA CCTCCCTAC CTGGTAGCGT ACCAAGCCAC CGTGTGCGCT  
 CACCCCTCTT GGAAGGAATG GACCATCGCA TGGTCGGTG GCACACCGCA

+2 R A Q A P P P S W D Q M W K C L I  
 3701 AGGGCTCAAG CCCCTCCCCC ATCGTGGGAC CAGATGTGGA AGTGTGGAT  
 TCCCGAGTTC GGGGAGGGGG TAGCACCCCTG GTCTACACCT TCACAAACTA

+2 R L K P T L H G P T P L L Y R L  
 3751 TCGCCTCAAG CCCACCTCC ATGGGCAAC ACCCTGCTA TACAGACTGG  
 AGCGGAGTTC GGGTGGGAGG TACCCGGTTG TGGGACGAT ATGTCTGACC

+2 G A V Q N E I T L T H P V T K Y I  
 3801 GCGCTGTTCA GAATGAAATC ACCCTGACGC ACCCAGTCAC CAAATACATC  
 CGCGACAAGT CTTACTTTAG TGGGACTGCG TGGGTCACTG GTTTATGTAG

+2 M T C M S A D L E V V T S T W V L  
 3851 ATGACATGCA TGTCGGCCGA CCTGGAGGTC GTCACCGAGCA CCTGGGTGCT  
 TACTGTACGT ACAGCCGGCT GGACCTCCAG CAGTGTGCGT GGACCCACGA

+2 V G G V L A A L A A Y C L S T G  
 3901 CGTTGGCGGC GTCTGGCTG CTTTGGCCGC GTATTGCGCTG TCAACAGGCT  
 GCAACCGCCG CAGGACCGAC GAAACCGGGCG CATAACGGAC AGTTGTCCGA

pCMV-NS34A  
FIGURE 9 - Page 6

+2 C V V I V G R V V L S G K P A I I  
3951 GCGTGGTCAT AGTGGGCAGG GTCGTCTTGT CCGGGAAAGCC GGCAATCATA  
CGCACCAAGTA TCACCCGTCC CAGCAGAACAA GGCCCTTCGG CCGTTAGTAT

+2 P D R E V L Y R E F D E M E E C  
4001 CCTGACAGGG AAGTCCTCTA CCGAGAGTTC GATGAGATGG AAGAGTGCTA  
GGACTGTCCC TTCAGGAGAT GGCTCTCAAG CTACTCTACC TTCTCACGAT

BamHI MluI

4051 GCGATCCACTA CGCGTTAGAG CTCGCTGATC AGCCTCGACT GTGCCTCTA  
CCTAGGTGAT GCGCAATCTC GAGCGACTAG TCGGAGCTGA CACGGAAGAT

4101 GTTGCCAGCC ATCTGTTGTT TGCCCCCTCCC CCGTGCCTTC CTTGACCCCTG  
CAACGGTCGG TAGACAACAA ACGGGGAGGG GGCACGGAAG GAACTGGGAC

4151 GAAGGTGCCA CTCCCCACTGT CCTTCCCTAA TAAAATGAGG AAATTGCATC  
CTTCCACGGT GAGGGTGACA GGAAAGGATT ATTTTACTCC TTTAACGTAG

4201 GCATTGTCTG AGTAGGTGTC ATTCTATTCT GGGGGGTGGG GTGGGGCAGG  
CGTAACAGAC TCATCCACAG TAAGATAAGA CCCCCCACCC CACCCCGTCC

4251 ACAGCAAGGG GGAGGATTGG GAAGACAATA GCAGGCATGC TGGGGAGCTC  
TGTGTTCCC CCTCCTAACCC CTTCTGTTAT CGTCCGTACG ACCCCCTCGAG

4301 TTCCGCTTCC TCGCTCACTG ACTCGCTGCG CTCGGTCTT CGGCTGCGGC  
AAGGCGAAGG AGCGAGTGAC TGAGCGACGC GAGCCAGCAA GCCGACGCCG

4351 GAGCGGTATC AGCTCACTCA AAGGCGGTAA TACGGTTATC CACAGAATCA  
CTCGCCATAG TCGAGTGAGT TTCCGCCATT ATGCCAATAG GTGTCTTAGT

4401 GGGGATAAACG CAGGAAAGAA CATGTGAGCA AAAGGCCAGC AAAAGGCCAG  
CCCCTATTGC GTCCTTCTT GTACACTCGT TTTCCGGTCG TTTTCCGGTC

4451 GAACCGTAAA AAGGCCCGT TGCTGGCGTT TTTCCATAGG CTCCGCCCG  
CTTGGCATT TTCCGGCGCA ACGACCGCAA AAAGGTATCC GAGGCGGGGGG

4501 CTGACGAGCA TCACAAAAAT CGACGCTCAA GTCAGAGGTG GCGAAACCCG  
GACTGCTCGT AGTGTTTTA GCTGCGAGTT CAGTCTCCAC CGCTTGGC

4551 ACAGGACTAT AAAGATACCA GGCCTTCTCC CCTGGAAAGCT CCCTCGTGC  
TGTCTGATA TTTCTATGGT CCGCAAAGGG GGACCTTCGA GGGAGCACGC

4601 CTCTCCTGTT CCGACCCCTGC CGCTTACCGG ATACCTGTCC GCCTTCTCC  
GAGAGGACAA GGCTGGGACG GCGAATGGCC TATGGACAGG CGGAAAGAGG

4651 CTTCGGGAAG CGTGGCGCTT TCTCAATGCT CACGCTGTAG GTATCTCAGT  
GAAGCCCTTC GCACCGCGAA AGAGTTACGA GTGCGACATC CATAGAGTCA

4701 TCGGTGAGG TCGTTCGCTC CAAGCTGGC TGTTGTGACG AACCCCCCGT  
AGCCACATCC AGCAAGCGAG GTTCGACCCG ACACACGTGC TTGGGGGGCA

4751 TCAGCCCCGAC CGCTGGCGCT TATCCGGTAA CTATCGTCTT GAGTCCAACC  
AGTCGGGCTG GCGACGCGGA ATAGGCCATT GATAGCAGAA CTCAGGTTGG

4801 CGGTAAGACA CGACTTATCG CCACCTGGCAG CAGCCACTGG TAACAGGATT  
GCCATTCTGT GCTGAATAGC GGTGACCGTC GTCGGTGACC ATTGTCTAA

## FIGURE 9 - Page 7

4851 AGCAGAGCGA GGTATGTAGG CGGTGCTACA GAGTTCTTGA AGTGGTGGCC  
 TCGTCTCGCT CCATACATCC GCCACGATGT CTCAAGAACT TCACCACCGG  
 4901 TAACTACGGC TACACTAGAA GGACAGTATT TGGTATCTGC GCTCTGCTGA  
 ATTGATGCCG ATGTGATCTT CCTGTCATAA ACCATAGACG CGAGACGACT  
 4951 AGCCAGTTAC CTTCGGAAAA AGAGTTGGTA GCTCTTGATC CGGCAAACAA  
 TCGGTCAATG GAAGCCTTT TCTCAACCAT CGAGAACTAG GCGTTTGTT  
 5001 ACCACCGCTG GTAGCGGTGG TTTTTTGTT TGCAAGCAGC AGATTACGCG  
 TGGTGGCGAC CATGCCACC AAAAAAAACAA ACAGTCGTCG TCTAATGCGC  
 5051 CAGAAAAAAA GGATCTCAAG AAGATCCTT GATCTTTCT ACAGGGTCTG  
 GTCTTTTTT CCTAGAGTTC TTCTAGGAAA CTAGAAAAGA TGCCCCAGAC  
 5101 ACGCTCAGTG GAACGAAAAC TCACGTTAAG GGATTTGGT CATGAGATTA  
 TCGAGTCAC CTTGCTTTG AGTCAATT CCTAAAACCA GTACTCTAAT  
 5151 TCAAAAAGGA TCTTCACCTA GATCCTTTA AATTAAAAT GAAGTTTAA  
 AGTTTTCT AGAAGTGGAT CTAGGAAAAT TTAATTTTA CTTCAAAATT  
 5201 ATCAATCTAA AGTATATATG AGTAAACTTG GTCTGACAGT TACCAATGCT  
 TAGTTAGATT TCATATATAC TCATTGAAAC CAGACTGTCA ATGGTTACGA  
 5251 TAATCACTGA GGCACCTATC TCAGCGATCT GTCTATTTCG TTCATCCATA  
 ATTAGTCACT CCGTGGATAG AGTCGCTAGA CAGATAAAAGC AAGTAGGTAT  
 5301 GTTGCCTGAC TCCCCGTCGT GTAGATAACT ACGATAACGGG AGGGCTTACC  
 CAACGGACTG AGGGGCAGCA CATCTATTGA TGCTATGCC CCCCCAATGG  
 5351 ATCTGGCCCC AGTGCTGCAA TGATACCGCG AGACCCACGC TCACCGGCTC  
 TAGACCGGGG TCACGACGTT ACTATGGCGC TCTGGGTGCG AGTGGCCGAG  
 5401 CAGATTTATC AGCAATAAAC CAGCCAGCCG GAAGGGCCGA GCGCAGAACT  
 GTCTAAATAG TCGTTATTTG GTCGGTCGGC CTTCCCGGT CGCGTCTTCA  
 5451 GGTCCCTGCAA CTTTATCCGC CTCCCATCCAG TCTATTAATT GTTGCCTGGGA  
 CCAGGACGTT GAAATAGGCG GAGGTAGGTC AGATAATTAA CAACGGCCCT  
 5501 AGCTAGAGTA AGTAGTTCGC CAGTTAATAG TTTGCGCAAC GTTGTGCCA  
 TCGATCTCAT TCATCAAGCG GTCAATTATC AAACCGCGTT CAACAACCGT  
 5551 TTGCTACAGG CATCGTGGTG TCACGCTCGT CGTTGGTAT GGCTTCATTG  
 AACGATGTCC GTAGCACCAC AGTGCAGAGCA GCAAACCATCA CCGAAGTAAG  
 5601 AGCTCCGGTT CCCAACGATC AAGGCGAGTT ACATGATCCC CCATGTTGTG  
 TCGAGGCCAA GGGTTGCTAG TTCCGCTCAA TGTACTAGGG GGTACAACAC  
 5651 CAAAAAAAGCG GTTAGCTCCT TCGGTCTCC GATCGTGTC AGAAGTAAGT  
 GTTTTTCCG CAATCGAGGA AGCCAGGAGG CTAGCAACAG TCTTCATTCA  
 5701 TGGCCGCAGT GTTATCACTC ATGGTTATGG CAGCACTGCA TAATTCTCTT  
 ACCGGCGTCA CAATAGTGAG TACCAATACC GTCGTGACGT ATTAAGAGAA  
 5751 ACTGTCATGC CATCGTAAG ATGTTTTCT GTGACTGGTG AGTACTCAAC  
 TGACAGTACG GTAGGCATTC TACGAAAAGA CACTGACCAAC TCATGAGTTG

**FIGURE 9 - Page 8**

5801 CAAGTCATTC TGAGAATAGT GTATGCGCG ACCGAGTTGC TCTTGCCCCGG  
GTTCAAG ACTCTTATCA CATA CGCCGC TGGCTCAACG AGAACGGGCC

5851 CGTCAATACG GGATAATACC GCGCCACATA GCAGAACTTT AAAAGTGCTC  
GCAGTTATGC CCTATTATGG CGCGGTGTAT CGTCTTGAAA TTTTCACGAG

5901 ATCATTGGAA AACGTTCTTC GGGGCGAAAA CTCTCAAGGA TCTTACCGCT  
TAGTAACCTT TTGCAAGAAG CCCCCTTT GAGAGTTCCCT AGAATGGCGA

5951 GTTGAGATCC AGTTCGATGT AACCCACTCG TGCACCCAAAC TGATCTTCAG  
CAACTCTAGG TCAAGCTACA TTGGGTGAGC ACGTGGGTG ACTAGAAGTC

6001 CATCTTTAC TTTCACCAGC GTTCTGGGT GAGCAAAAAC AGGAAGGCAA  
GTAGAAAATG AAAGTGGTCG CAAAGACCCA CTCGTTTTG TCCTCCGTT

6051 AATGCCGCAA AAAAGGGAAT AAGGGCGACA CGGAAATGTT GAATACTCAT  
TTACGGCGTT TTTCCCTTA TTCCCGCTGT GCCTTACAA CTTATGAGTA

6101 ACTCTTCCTT TTCAATATT ATTGAAGCAT TTATCAGGGT TATTGTCTCA  
TGAGAAGGAA AAAGTTATAA TAACTCGTA AATAGTCCCA ATAACAGAGT

6151 TGAGCGGATA CATATTGAA TGTATTTAGA AAAATAAACAA AATAGGGGTT  
ACTCGCCTAT GTATAAACTT ACATAAAATCT TTTTATTGTT TTATCCCCAA

6201 CCGCGCACAT TTCCCCGAAA AGTGCCACCT GACGTCTAAG AAACCATTAT  
GGCGCGTGTAAAGGGCTTT TCACGGTGGA CTGCAGATTC TTTGGTAATA

6251 TATCATGACA TTAACCTATA AAAATAGGCG TATCACGAGG CCCTTCGTC  
ATAGTACTGT AATTGGATAT TTTTATCCGC ATAGTGCCTCC GGGAAAGCAG

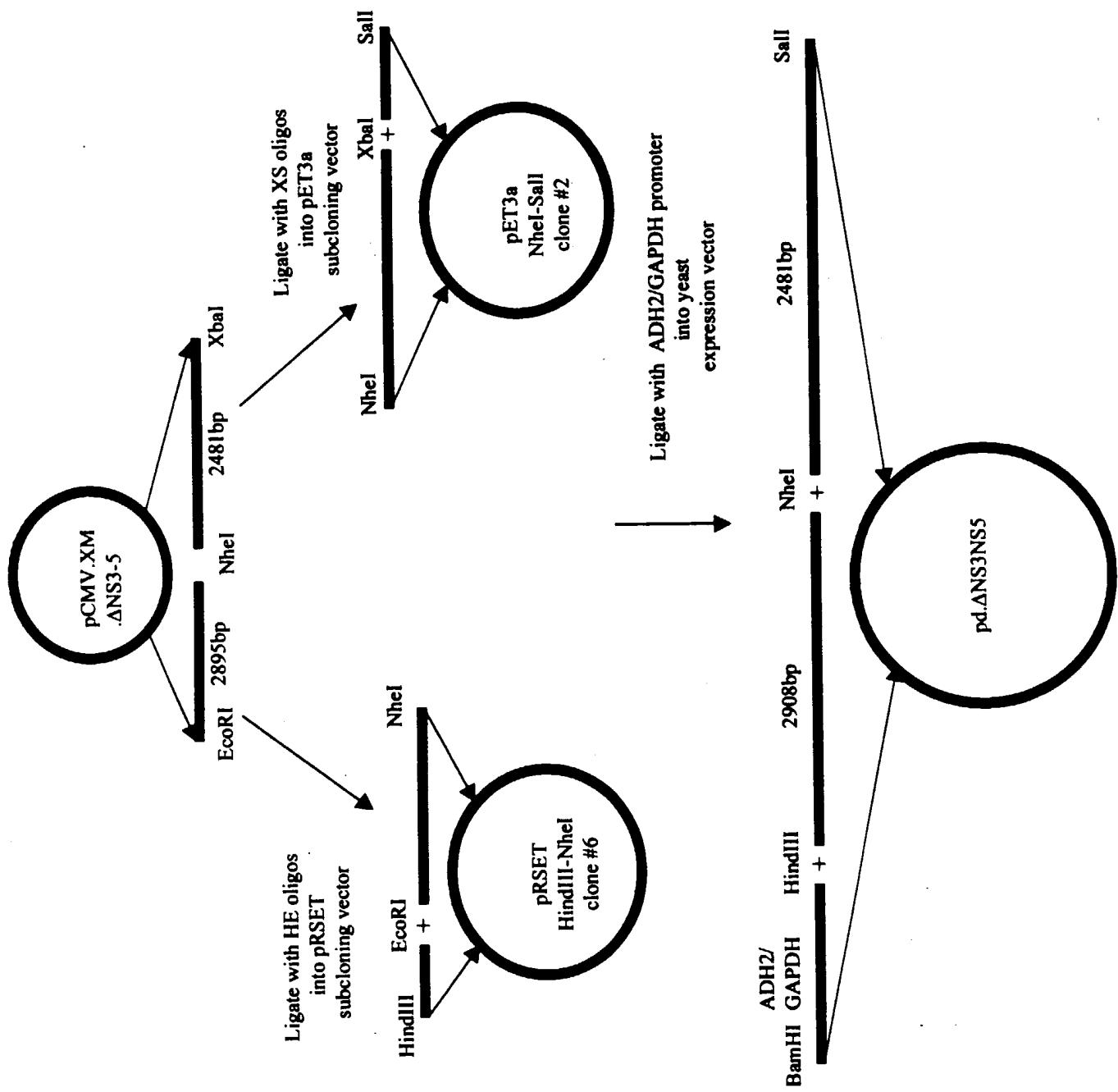


FIGURE 10

## FIGURE 11 - Page 1

Met Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val Leu Val  
 2 AGCTTACAAAACAAATTCCACCATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTA  
 TCGAATGTTTGTAAAGTGGTACCGACGTATACGTCGAGTCCCGATATTCCACGATCAT  
 ^ ^ ^  
 1 HIND3, 21 NCOI, 30 NDEI, 58 SCAI,  
  
 Leu Asn Pro Ser Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr Met Ser Lys Ala His Gly  
 62 CTCAACCCCTCTGTTGCTGCAACACTGGGCTTGCTTACATGTCCAAGGCTCATGGG  
 GAGTTGGGAGACAAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCC  
  
 Ile Asp Pro Asn Ile Arg Thr Gly Val Arg Thr Ile Thr Thr Gly Ser Pro Ile Thr Tyr  
 122 ATCGATCCTAACATCAGGACCGGGGTGAGAACAAATTACCAACTGGCAGCCCCATCACGTAC  
 TAGCTAGGATTGTAGTCCTGGCCCCACTCTGTAAATGGTGACCGTGGGTAGTGCATG  
 ^  
 122 CLAI,  
  
 Ser Thr Tyr Gly Lys Phe Leu Ala Asp Gly Gly Cys Ser Gly Gly Ala Tyr Asp Ile Ile  
 182 TCCACCTACGGCAAGTCCCTGCCGACGGCGGGTGCCTCGGGGGCGCTTATGACATAATA  
 AGGTGGATGCCGTTCAAGGAACGGCTGCCGCCACGAGCCCCCGCGAATACTGTATTAT  
  
 Ile Cys Asp Glu Cys His Ser Thr Asp Ala Thr Ser Ile Leu Gly Ile Gly Thr Val Leu  
 242 ATTTGTGACGAGTGCCTACCGGATGCCACATCCATCTGGGCATTGGCACTGTCCTT  
 TAAACACTGCTCACGGTGAGGTGCCTACGGTAGGTAGAACCCGTAACCGTGACAGGAA  
  
 Asp Gln Ala Glu Thr Ala Gly Ala Arg Leu Val Val Leu Ala Thr Ala Thr Pro Pro Gly  
 302 GACCAAGCAGAGACTGCAGGGCGAGACTGGTTGTGCCTGCCACCGCCACCCCTCCGGGC  
 CTGGTTCGTCTGACGCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGAGGCCCG  
 ^  
 309 ALWN1,  
  
 Ser Val Thr Val Pro His Pro Asn Ile Glu Glu Val Ala Leu Ser Thr Thr Gly Glu Ile  
 362 TCCGTCACTGTGCCCATCCAACATCGAGGAGGTGCTCTGCCACCACGGAGAGATC  
 AGGCAGTGACACGGGGTAGGGTTGTAGCTCCTCCAACCGAGACAGGTGGTGGCCTCTAG  
  
 Pro Phe Tyr Gly Lys Ala Ile Pro Leu Glu Val Ile Lys Gly Gly Arg His Leu Ile Phe  
 422 CCTTTTACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGAGACATCTCATCTTC  
 GGAAAAATGCCGTTCCGATAGGGGAGCTCATTAGTCCCCCCTCTGTAGAGTAGAAG  
  
 Cys His Ser Lys Lys Cys Asp Glu Leu Ala Ala Lys Leu Val Ala Leu Gly Ile Asn  
 482 TGTCAATTCAAAGAAGAAGTGCACGAACTCGCCGAAAGCTGGTCGCATTGGCATTCAAT  
 ACAGTAAGTTCTCTCACGCTGCTTGAGCGGGTTTCGACCAGCGTAACCGTAGTTA  
  
 Ala Val Ala Tyr Tyr Arg Gly Leu Asp Val Ser Val Ile Pro Thr Ser Gly Asp Val Val  
 542 GCGCGGCTACTACCGCGGTCTGACGTGTCCCGTCATCCCGACCAGCGCGATGGTGC  
 CGGCACCGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCCGCTACAAACAG  
 ^ ^  
 556 SAC2, 566 DRD1,  
  
 Val Val Ala Thr Asp Ala Leu Met Thr Gly Tyr Thr Gly Asp Phe Asp Ser Val Ile Asp  
 602 GTCGTGGCAACCGATGCCCTCATGACCGGCTATAACCGGCACCTCGACTCGGTGATAGAC  
 CAGCACCGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTG  
 ^  
 621 BSPH1,  
  
 Cys Asn Thr Cys Val Thr Gln Thr Val Asp Phe Ser Leu Asp Pro Thr Phe Thr Ile Glu

## FIGURE 11 - Page 2

662 TGCAATACGTGTCAACCCAGACAGTCGATTCAGCCTTGACCCCTACCTTCACCATTGAG  
 ACGTTATGCACACAGTGGTCTGTCAGCTAAAGTCGGAACGGGATGGAAGTGGTAAC  
 ThrIleThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArg  
 722 ACAATCACGCTCCCCAAGATGCTGTCAGCTAAACGTCGGGGCAGGACTGGCAGG  
 TGTTAGTGCAGGGGGTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCTGACCGTCC  
 GlyLysProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAsp  
 782 GGGAAAGCCAGGCATCTACAGATTGTGGCACCGGGGAGCGCCCTCCGGCATGTTGAC  
 CCCTTCGGTCCGTAGATGTCTAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTG  
 822 BGLI, 839 DRD1,  
 SerSerValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAla  
 842 TCGTCCGTCCTCTGTGAGTGTATGACGCAAGCTGTGCTTGGTATGAGCTCACGCCGCC  
 AGCAGGCAGGAGACACTCACGATACTGCGTCCGACACGAAACCATACTCGAGTGCGGGCGG  
 887 SACI,  
 GluThrThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAsp  
 902 GAGACTACAGTTAGGCTACGAGCGTACATGAACACCCGGGGCTTCCGTGTGCCAGGAC  
 CTCTGATGTCAATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCCTG  
 937 SMAI XMAI,  
 HisLeuGluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeu  
 962 CATCTTGAATTTGGGAGGGCGTCTTACAGGCCTCACTCATATAGATGCCACTTCTA  
 GTAGAACTTAAACCCCTCCCGAGAAATGTCCGGAGTGAATATCTACGGGTGAAAGAT  
 991 STUI,  
 SerGlnThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrVal  
 1022 TCCCAGACAAAGCAGGTGGGAGAACCTTCTTACCTGGTAGCGTACCAAGCCACCGTG  
 AGGGTCTGTTCTCACCCCTCTTGAAGGAATGGACCATCGCATGGTCGGTGGCAC  
 1075 DRA3,  
 CysAlaArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArg  
 1082 TCGCCTAGGGCTCAAGCCCCCCCCATCGTGGGACCAGATGTGGAAGTGGATTGATTG  
 ACGCGATCCCGAGTTGGGGAGGGGTAGCACCCCTGGTCTACACCTTCACAAACTAACGCG  
 LeuLysProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsn  
 1142 CTCAGCCCCACCCCTCCATGGGCCAACACCCCTGCTATACAGACTGGCGCTGTTCAAGAAT  
 GAGTTGGGTGGAGGTACCCGGTTGTGGGACGATATGTCTGACCCGGACAAAGTCTTA  
 1156 NCOI,  
 GluIleThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeu  
 1202 GAAATCACCCCTGACGCACCCAGTCACCAAATACATGACATGCGATGTCGGCGACCTG  
 CTTTAGTGGGACTGCGTGGGTCACTGGTTATGTAGTACTGTACGTACAGCCGGCTGGAC  
 1236 BSPH1, 1240 DRD1, 1243 AVA3, 1251 EAG1 XMA3, 1256 DRD1,  
 GluValValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyr  
 1262 GAGGTGCGTCACGAGCACCTGGGTGCTGTTGGCGCGTCTGGCTGCTTGGCCGCGTAT  
 CTCCAGCAGTGCCTGGACCCACGAGCAACCGCCGCAGGACCGACGAAACCGGGCGCAT

## 8

### FIGURE 11 - Page 3

1322 CysLeuSerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAla  
 TGCCTGTCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCCGGAAAGCCGGCA  
 ACGGACAGTTGTCCGACGCACCAGTATCACCCGTCCCAGCAGAACAGGCCCTCGGCCGT  
 ^  
 1375 NAEI,  
 IleIleProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGln  
 1382 ATCATAACCTGACAGGGAAAGTCCTCTACCGAGAGTTCGATGAGATGGAAGAGTGCTCTCAG  
 TAGTATGGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTTACCTTCACGAGAGTC  
 ^  
 1391 DRD1,  
 HisLeuProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeu  
 1442 CACTTACCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAACGCCCTC  
 GTGAATGGCATGTAGCTCGTCCCTACTACGAGCGGCTCGTCAAGTTCGTCTCCGGAG  
 GlyLeuLeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsn  
 1502 GGCCTCCTGCAGACCGCGTCCCGTCAGGCAGAGTTATGCCCTGCTGCCAGACCAAC  
 CCGGAGGACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTG  
 ^ ^  
 1508 PSTI, 1513 TTH3I,  
 TrpGlnLysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGln  
 1562 TGGCAAAACTCGAGACCTCTGGCGAAGCATAATGTGGAACCTTCATCAGTGGGATACAA  
 ACCGTTTTGAGCTCTGGAAAGACCCGCTCGTACACCTGAAGTAGTCACCCATATGTT  
 ^ ^  
 1571 XHOI, 1592 NDEI,  
 TyrLeuAlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPhe  
 1622 TACTTGGCGGGCTTGTCAACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTT  
 ATGAACCGCCCCAACAGTTGCGACGGACCATTGGGGCGTAACGAAGTAACCGAAAAA  
 ^  
 1649 BSTE2,  
 ThrAlaAlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGly  
 1682 ACAGCTGCTGTCACCAGCCCACTAACCAACTAGCCAAACCTCTTCAACATATTGGGG  
 TGTCGACGACAGTGGTCGGGTGATTGGTATCGGTTGGGAGGAGATTGTATAACCCC  
 ^  
 1683 ALWN1 PVU2,  
 GlyTrpValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGly  
 1742 GGGTGGGTGGCTGCCAGCTCGCCGCCCGGTGCCGACTGCCTTGTCGGCGCTGGC  
 CCCACCCACCGACGGGTCGAGCGGGCCACGGCGATGACGGAAACACCCCGCGACCG  
 ^  
 1800 ESP1,  
 LeuAlaGlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAla  
 1802 TTAGCTGGCGCCCATCGGCAGTGTGGACTGGGAAGGTCTCATAGACATCCTGCA  
 AATCGACCGCGGGTAGCCGTACAACCTGACCCCTCAGGAGTATCTGTAGGAACGT  
 ^  
 1808 KAS1 NARI,  
 GlyTyrGlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluVal  
 1862 GGGTATGGCGCGGGCGTGGCGGGAGCTTGTGGCATTCAAGATCATGAGCGGTGAGGTC  
 CCCATACCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAG  
 ^

## FIGURE 11 - Page 4

1884 SACI, 1905 BSPH1,

1922 ProSerThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuVal  
 CCCTCCACGGAGGACCTGGTCAATCTACTGCCGCCATCCTCTGCCGGAGCCCTCGTA  
 GGGAGGTGCCTCTGGACCAGTTAGATGACGGCGGTAGGAGAGCAGGGCCTCGGGAGCAT

1934 TTH3I,

1982 ValGlyValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaVal  
 GTCGGCGTGGTCTGTGCAGCAATACTGCCGCCACGTTGGCCCGAGGGGGCAGTG  
 CAGCCGACACAGACGTGTTATGACGCCGTGCAACCAGGGCCGCTCCCCGTAC

2010 NAEI, 2023 SMAI XMAI,

2042 GlnTrpMetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHis  
 CAGTGGATGAACCGGCTGATAGCCTCGCCTCCGGGGAACATGTTCCCCCACGCAC  
 GTCACCTACTTGGCCGACTATCGGAAGCGGAGGGCCCGCTGGTACAAAGGGGGTGGCGT

2073 SMAI XMAI, 2099 DRA3,

2102 TyrValProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrVal  
 TACGTGCCGGAGAGCGATGCAGCTGCCCGCTACTGCCATACTCAGCAGCCTCACTGTA  
 ATGCACGGCCTTCGCTACGTGACGGCGCAGTGACGGTATGAGTCGTCGGAGTGACAT

2121 PVU2,

2162 ThrGlnLeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSer  
 ACCCAGCTCCTGAGGCGACTGCACCACTGGATAAGCTCGGAGTGTACCACTCCATGCTCC  
 TGGGTGAGGACTCCGCTGACGTGGTACCTATTGAGCCTACATGGTGAGGTACGAGG

2165 ALWN1, 2170 MST2,

2222 GlySerTrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThr  
 GGTTCCTGGCTAAGGGACATCTGGACTGGATATGCGAGGTGTTGAGCGACTTTAACACC  
 CCAAGGACCGATTCCCTGTAGACCTACGCTCCACAACTCGCTGAAATTCTGG

2226 ECON1,

2282 TrpLeuLysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArg  
 TGGCTAAAAGCTAACGCTCATGCCACAGCTGCCCTGGGATCCCCTTGTGTCCTGCCAGCGC  
 ACCGATTTGATTCGAGTACGGTGTGACGGACCCCTAGGGGAAACACAGGACGGTCGCG

2291 ESP1, 2306 PVU2, 2316 BAMHI,

2342 GlyTyrLysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAla  
 GGGTATAAGGGGGCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCT  
 CCCATATTCCCCAGACCGCTCCCTGCCGTAGTACGTGTGAGCGACGGTACACCTCGA

2402 GluIleThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArg  
 GAGATCACTGGACATGTCAAAACGGGACGATGAGGATCGTCGGTCTAGGACCTGCAGG  
 CTCTAGTGACCTGTACAGTTTGCCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCC

2431 BSAB1, 2447 AVR2, 2454 SSE83871, 2455 PSTI,

2462 AsnMetTrpSerGlyThrPheProIleAsnAlaTyrThrGlyProCysThrProLeu  
 AACATGTGGAGTGGGACCTCCCCATTAAATGCCTACACCACGGGCCCTGTACCCCCCTT  
 TTGTACACCTCACCCCTGGAAGGGTAATTACGGATGTGGTGCCGGGGACATGGGGGAA

**FIGURE 11 - Page 5**

2486 ASE1, 2503 APA1,

2522 ProAlaProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIle  
CCTGCGCCGAACTACACGTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATA  
GGACGCGGTTGATGTGCAAGCGCGATAACCTCCCACAGACGTCTCCTTATGCACCTCTAT

2559 PST1,

2582 ArgGlnValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysPro  
AGGCAGGTGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCG  
TCCGTCCACCCCCCTGAAGGTGATGCACTGCCATACTGATGACTGTTAGAATTACGGGC

2600 DRA3,

2642 CysGlnValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPhe  
TGCCAGGTCCCATCGCCCGAATTTCACAGAATTGGACGGGTGCGCCTACATAGGTTT  
ACGGTCCAGGGTAGCGGGCTAAAAAGTGTCTAACCTGCCAACGCGGATGTATCCAAA

2702 AlaProProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGlu  
GCGCCCCCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTAGAGTAGGACTCCACGAA  
CGCGGGGGACGTTGGAACGACGCCCTCCATAGTAAGTCTCATCCTGAGGTGCTT

2762 TyrProValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSer  
TACCCGGTAGGGTCGCAATTACCTTGCAGCCCCGAACCGGACGTGGCCGTGTTGACGTCC  
ATGGGCCATCCCAGCGTTAATGGAACGCTGGCCTGCACCGGACAAACTGCAGG

2763 HGIE2, 2815 AAT2,

2822 MetLeuThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGly  
ATGCTCACTGATCCCTCCCATATAAACAGCAGAGGGCGCCGGCGAAGGTTGGCGAGGGGA  
TACGAGTGAATAGGGAGGGTATATTGTCGTCTCCGCCGGCCCTCCAACCGCTCCCT

2856 EAG1 XMA3,

2882 SerProProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAla  
TCACCCCCCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCGCTCCATCTCAAGGCA  
AGTGGGGGAGACACCGGTCGAGGAGCCATCGGTCGATAGGCGAGGTAGAGAGTCCGT

2895 BALI, 2909 NHEI,

2942 ThrCysThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrp  
ACTTGACCGCTAACCATGACTCCCCGTGATGCTGAGCTCATAGAGGCCAACCTCCTATGG  
TGAACGTGGCGATTGGTACTGAGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACC

2972 ESP1, 2975 SAC1,

3002 ArgGlnGluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeu  
AGGCAGGAGATGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTG  
TCCGTCCCTACCCGCCGTGTTAGTGGCTCCAACTCAGTCTTTGTTCACCAACTAAGAC

3062 AspSerPheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGlu  
GACTCCTCGATCCGTTGTGGCGGAGGAGGACGAGCAGGGAGATCTCCGTACCCGCAGAA  
CTGAGGAAGCTAGGCGAACACCGCCTCCCTGCTGCCCTAGAGGCATGGCGTCTT

3102 BGL2,

## FIGURE 11 - Page 6

3122 IleLeuArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyr  
 ATCCTGCGGAAGTCTCGGAGATTGCCAGGCCCTGCCCGTTGGCGCGCCGGACTAT  
 TAGGACGCCCTCAGAGCCTCTAACGGTCCGGACGGCAAACCCGCGCCGGCTGATA  
 ^ ^  
 3149 ALWN1, 3170 EAG1 XMA3,

3182 AsnProProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGly  
 AACCCCCCGCTAGTGGAGACGTGGAAAAGCCGACTACGAACCACCTGTGGTCCATGGC  
 TTGGGGGGCGATCACCTCTGCACCTTCAGGGCTGATGCTGGTGGACACCAGGTACCG  
 ^ ^  
 3223 HGIE2, 3235 NCOI,

3242 CysProLeuProProLysSerProProValProProProArgLysLysArgThrVal  
 TGCCCCTTCCACCTCAAAGTCCCCTCTGTGCCTCCGCTCGGAAGAACGGACGGTG  
 ACGGCGAAGGTGGAGGTTCAAGGGAGGACACGGAGGCGAGCCTCTCGCCTGCCAC

3302 ValLeuThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGly  
 GTCCTCACTGAATCAACCTATCTACTGCCTTGCGAGCTGCCACCAGAAGCTTGGC  
 CAGGAGTGACTTAGTGGATAGATGACGGAACCGCTCGAGCGGTGGCTTCGAAACCG  
 ^ ^  
 3338 SACI, 3352 HIND3,

3362 SerSerSerThrSerGlyIleThrGlyAsnThrThrSerSerGluProAlaPro  
 AGCTCCTCAACTCCGGCATTACGGCGACAATACGACAACATCCTCTGAGCCCGCCCCT  
 TCGAGGAGTTGAAGGCCGTAATGCCGCTGTTATGCTGTTAGGAGACTCGGGCGGGGA

3422 SerGlyCysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGly  
 TCTGGCTGCCCGGCGACTCCGACGCTGAGTCCTATTCCCTCATGCCCGCTGGAGGGG  
 AGACCGACGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGACCTCCCC  
 ^  
 3443 EAM11051,

3482 GluProGlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsn  
 GAGCCTGGGATCCGGATCTTAGCGACGGTCATGGTCAACGGTCAGTAGTGAGGCCAAC  
 CTCGGACCCCTAGGCCTAGAATCGCTGCCAGTACCAGTTGCCAGTCATCACTCCGGTTG  
 ^ ^  
 3490 BAMHI, 3491 BSAB1, 3493 BSPE1,

3542 AlaGluAspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrPro  
 GCGGAGGATGTCGTGCTGCTCAATGTCTACTTGGACAGGCGCACTCGTCACCCCG  
 CGCCTCCTACAGCACACGAGTTACAGAATGAGAACCTGCCGCTGAGCAGTGGGGC  
 ^  
 3595 DRA3,

3602 CysAlaAlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHis  
 TGCGCCGCGGAAGAACAGAAACTGCCCATCAATGCACTAACGAACTCGTTGCTACGTCAC  
 ACGGCGCCCTTGTCTTGACGGTAGTTACGTGATTGAGCAACGATGCACTGAGTGT  
 ^ ^  
 3606 SAC2, 3617 ALWN1, 3661 PFLM1,

3662 HisAsnLeuValTyrSerThrSerArgSerAlaCysGlnArgGlnLysLysValThr  
 CACAATTGGTGTATTCCACCACCTCACGCAGTGCCTGCCAAAGGCAGAAGAAAGTCACA  
 GTGTTAACACACATAAGGTGGAGTGCCTCACGAACGGTTCCGTCTTCAGTGT  
 ^  
 3687 DRA3,

PheAspArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAla

## FIGURE 11 - Page 7

3722 TTTGACAGACTGCAAGTTCTGGACAGCATTACCAGGACGTACTCAAGGAGGTTAAAGCA  
 AAACTGTCTGACGTTCAAGACCTGTCGGAATGGTCCTGCATGAGTTCTCCAATTTCGT  
 AlaAlaSerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrPro  
 3782 GCGCGTCAAAAGTGAAGGCTAATTGCTATCCGTAGAGGAAGCTGCAGCCTGACGCC  
 CGCCGCAGTTTCACTCCGATTGAACGATAGGCATCTCCTCGAACGTCGGACTGCAGGG  
 ^  
 3822 HIND3,  
 ProHisSerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArg  
 3842 CCACACTCAGCCAAATCCAAGTTGGTATGGGGAAAAGACGTCGTTGCCATGCCAGA  
 GGTGTGAGTCGGTTAGGTTCAAACCAATACCCGTTCTGCAGGCAACGGTACGGTCT  
 ^ ^  
 3881 AAT2, 3896 BGLI,  
 LysAlaValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrPro  
 3902 AAGGCCGTAAACCCACATCAACTCCGTGAAAGACCTCTGGAAGACAATGTAACACCA  
 TTCCGGCATTGGGTGTAGTTGAGGCACACCTTCTGGAAGACCTCTGTTACATTGGGT  
 ^  
 IleAspThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGly  
 3962 ATAGACACTACCACATGGCTAAGAACGAGGTTCTGCGTCAGCCTGAGAAGGGGGGT  
 TATCTGTGATGGTAGTACCGATTCTGCTCCAAAAGACGCAAGTCGGACTCTCCCCCA  
 ^  
 ArgLysProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMet  
 4022 CGTAAGCCAGCTCGTCTCATCGTGTCCCCGATCTGGCGTGCCTGAGAAGGGGGATG  
 GCATTCGGTCGAGCAGAGTAGCACAAGGGCTAGACCCGCACGCGCACACGCTTTCTAC  
 ^  
 AlaLeuTyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPhe  
 4082 GCTTTGTACGACGTGGTACAAAGCTCCCCTGGCGTGTGGGAAGCTCCTACGGATTC  
 CGAAACATGCTGCACCAATGTTCGAGGGAAACCGGCACTACCCTCGAGGATGCCTAAG  
 ^  
 GlnTyrSerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysThr  
 4142 CAATACTACCAGGACAGCGGGTTGAATTCTCGCAAGCGTGGAAAGTCCAAGAAAACC  
 GTTATGAGTGGCCTGCGCCCAACTTAAGGAGCACGTCGACCTCAGGTTCTTGG  
 ^  
 4166 ECORI,  
 ProMetGlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIle  
 4202 CCAATGGGTTCTCGTATGATACCCGCTGCTTGACTCCACAGTCAGTGAGAGCGACATC  
 GGTTACCCCAAGAGCATACTATGGCGACGAAACTGAGGTGTCAGTGACTCTCGCTGTAG  
 ^ ^  
 4235 DRD1, 4242 ALWN1,  
 ArgThrGluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIle  
 4262 CGTACGGAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCCAAGCCCGGTGGCCATC  
 GCATGCCTCCTCCGTTAGATGGTTACAACACTGGAGCTGGAGCTGGGGTTCGGGCGCACCGGTAG  
 ^ ^  
 4307 BGLI, 4314 BALI,  
 LysSerLeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsn  
 4322 AAGTCCCTACCGAGAGGCTTATGTTGGGGCCCTCTTACCAATTCAAGGGGGAGAAC  
 TTCAGGGAGTGGCTCTCCGAAATACAACCCCGGGAGAATGGTTAAGTCCCCCTCTTG  
 ^  
 4351 APAI,  
 CysGlyTyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeu  
 4382 TGCAGGCTATCGCAGGTGCCGCGCAGCGCGTACTGACAACAGCTGTGGTAACACCCCTC

## FIGURE 11 - Page 8

ACGCCGATAGCGTCCACGGCGCGCTGCCGCATGACTGTTGATCGACACCATTGTGGGAG  
 4442 ThrCysTyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMet  
 ACTTGCTACATCAAGGCCGGCAGCGTGTGAGCCGCAGGGCTCCAGGACTGCACCATG  
 TGAACGATGTAGTCCGGGCCGTCGGACAGCTCGCGTCCCGAGGTCTGACGTGGTAC  
 4458 SMAI XMAI,  
 4502 LeuValCysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAla  
 CTCGTGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGTCCAGGAGGACGCG  
 GAGCACACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCTCCTGCGC  
 4514 DRD1, 4517 TTH3I,  
 4562 AlaSerLeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspPro  
 GCGAGCCTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCTGGGACCCC  
 CGCTCGGACTCTCGGAAGTGCCTCCGATACTGGTCATGAGGCGGGGGACCCCTGGG  
 4622 ProGlnProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAla  
 CCACAAACCAGAATACGACTTGGAGCTCATACATCATGCTCCTCCAACGTGTCAGTCGCC  
 GGTGTTGGTCTTATGCTAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGG  
 4643 SACI,  
 4682 HisAspGlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAla  
 CACGACGGCGCTGAAAGAGGGTCTACTACCTCACCCGTGACCCCTACAACCCCCCTCGCG  
 GTGCTGCCGCGACCTTCTCCAGATGATGGAGTGGACTGGATGTTGGGGAGCGC  
 4737 NRUI,  
 4742 ArgAlaAlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIle  
 AGAGCTCGTGGAGACAGCAAGACACACTCCAGTCATTCTGGCTAGGCAACATAATC  
 TCTCGACGCACCCCTCTGCGTTCTGTGAGGTCACTGTTAAAGGACCGATCCGTTATTAG  
 4802 MetPheAlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPheSerValLeu  
 ATGTTTGGCCCCACACTGTGGCGAGGATGATACTGATGACCCATTCTTAGCGTCCTT  
 TACAAACGGGGGTGTGACACCCGCTCCTACTATGACTACTGGTAAAGAAATCGCAGGAA  
 4812 PFLM1, 4813 DRA3,  
 4862 IleAlaArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSer  
 ATAGCCAGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGCCTGCTACTCC  
 TATCGGTCCTGGCGAACTGTCCGGAGCTAACGCTCTAGATGCCCGACGATGAGG  
 4899 BGL2,  
 4922 IleGluProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSer  
 ATAGAACCACTGGATCTACCTCCAATTCAGACTCCATGGCCTCAGCGCATTTC  
 TATCTGGTGACCTAGATGGAGGTTAGTAAGTTCTGAGGTACGGAGTCGCGTAAAGT  
 4960 NCOI,  
 4982 LeuHisSerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGly  
 CTCCACAGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGATGCCTCAGAAAATTGGG  
 GAGGTGTCAATGAGAGGTCCACTTAGTTATCCCACCGGCGTACGGAGTCTTTGAACCC  
 5021 SPHI, 5041 KPNI,

## FIGURE 11 - Page 9

5042 ValProProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAla  
GTACCGCCCTTGCAGCTTGGAGACACCGGGCCGGAGCGTCCCGCCTAGGCTTCTGGCC  
CATGGCGGGAACGCTCGAACCTCTGTGGCCCGGGCTCGCAGGCGCGATCCGAAGACCGG  
^  
5070 APAI, 5097 BALI,  
ArgGlyGlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLys  
5102 AGAGGAGGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAG  
TCTCCTCCGTCCCGACGGTATACACCGTTACGGAGAAGTTGACCGTCATTCTGTTTC  
^  
5119 NDEI,  
LeuLysLeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAla  
5162 CTCAAACTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTCACGGCT  
GAGTTGAGTGAGGTTATCGCCGGCGACCGGTCGACCTGAACAGGCCGACCAAGTGGCGA  
^ ^ ^  
5180 NOTI, 5181 EAG1 XMA3, 5188 BALI, 5192 PVU2,  
GlyTyrSerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrp  
5222 GGCTACAGCGGGGAGACATTATCACAGCGTGTCTCATGCCCGCCCCGCTGGATCTGG  
CCGATGTCGCCCCCTGTAAATAGTGTGCGCACAGAGTACGGGCCGGGCGACCTAGACC  
^  
5246 DRA3,  
PheCysLeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgOP  
5282 TTTTGCTACTCCTGCTGCTGCAGGGTAGGCATCTACCTCCTCCCCAACCGATGAAGG  
AAAACGGATGAGGACGAACGACGTCCCATCCGTAGATGGAGGAGGGTTGGCTACTTCC  
^  
5301 PSTI, 5331 HGIE2,  
5342 TTGGGGTAAACACTCCGGCTAAAAAAAAAAAAAAATCTAGAACCCGAGTCGAC  
AACCCCATTTGTGAGGCCGGATTTTTTTAGATCTGGGCTCAGCTG  
^ ^  
5378 XBAI, 5390 SALI,

FIGURE 12

→  
68  
Std  
PAS C.1 C.2

KDex

260

98

64

50

36

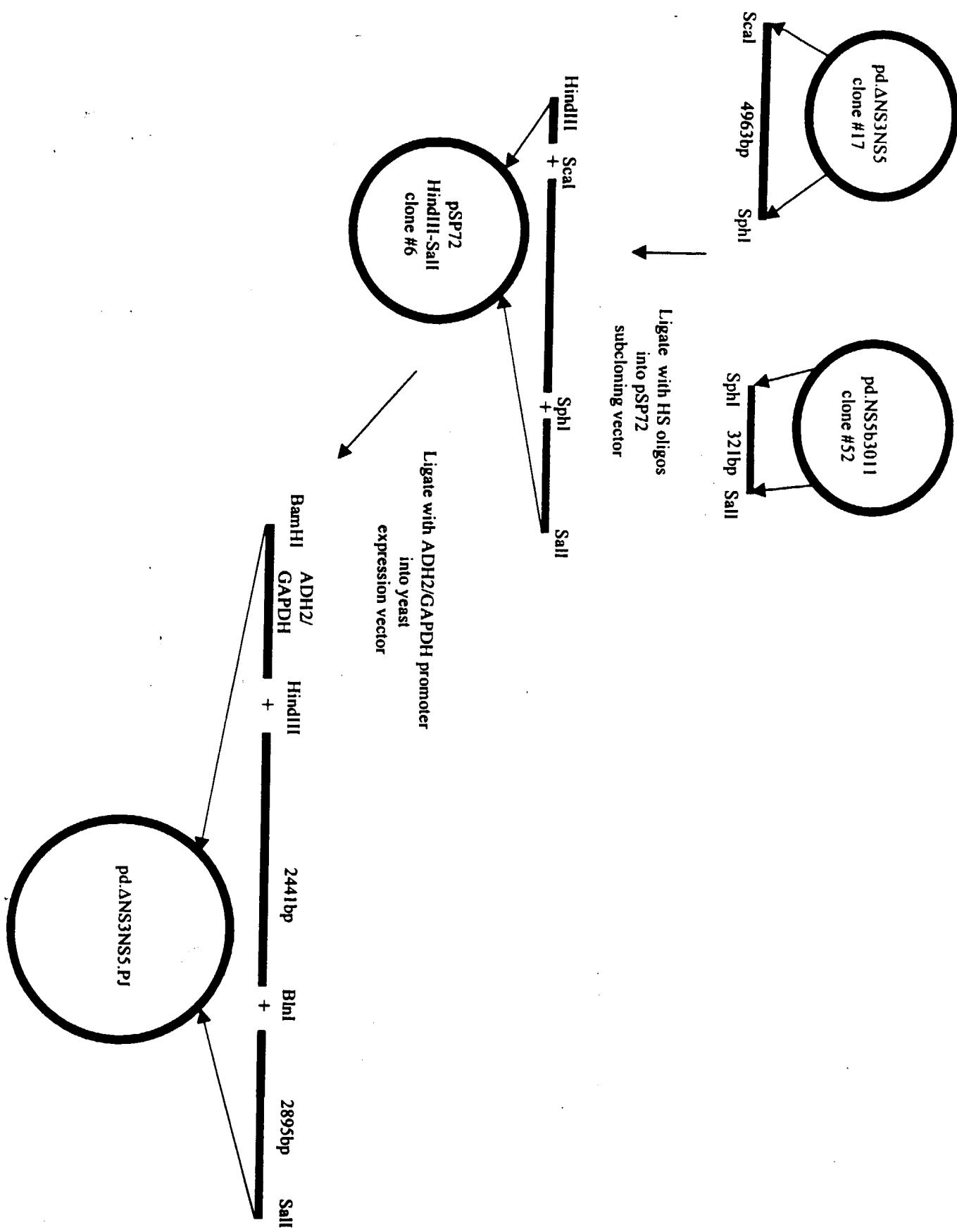
30

16

6

4

FIGURE 13



## FIGURE 14 - Page 1

MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn  
 2 AGCTTACAAAACAAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC  
 TCGAATGTTTGTACCGACGTATACGTCGAGTCCCAGTATTCCACGATCATGAGTTG  
 ^ ^ ^  
 1 HIND3, 24 NDEI, 52 SCAI,  
  
 ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp  
 62 CCCTCTGTTGCTGCAACACTGGCTTGGCTTACATGTCCAAGGCTCATGGGATCGAT  
 GGGAGACAAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCCTAGCTA  
 ^  
 116 CLAI,  
  
 ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr  
 122 CCTAACATCAGGACCGGGGTGAGAACAAATTACCACTGGCAGCCCCATCACGTACTCCACC  
 GGATTGTAGTCTGGCCCCACTCTTGTAAATGGTGACCGTCGGGTAGTGCATGAGGTGG  
  
 TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys  
 182 TACGGCAAGTTCTTGCCGACGGCGGGTGCTCGGGGGCGCTTATGACATAATAATTGT  
 ATGCCGTTCAAGGAACGGCTGCCGCCACGAGCCCCCGCGAATACTGTATTATTAAACA  
  
 AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln  
 242 GACGAGTGCCACTCCACGGATGCCACATCCATCTTGGCATTGGCACTGTCCTTGACCAA  
 CTGCTCACGGTGAGGTGCCTACGGTAGGTAGAACCCGTAACCGTGAAGGAACGGTT  
  
 AlaGluThrAlaGlyAlaArgLeuValLeuAlaThrAlaThrProProGlySerVal  
 302 GCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGCTCCGTC  
 CGTCTCTGACGCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCGAGGCAG  
 ^  
 303 ALWN1,  
  
 ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe  
 362 ACTGTGCCCATCCAAACATCGAGGAGGTTGCTCTGTCCACCACCGGAGAGATCCCTTT  
 TGACACGGGGTAGGGTTGTAGCTCCTCAAACGAGACAGGTGGCTCTAGGGAAAA  
  
 TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis  
 422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGAGACATCTCATCTTCTGT  
 ATGCCGTTCCGATAGGGGGAGCTTCATTAGTCCCCCCTCTGTAGAGTAGAACAGACTA  
  
 SerLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal  
 482 TCAAAGAAGTGCACGAACTCGCCGAAAGCTGGTCGATTGGCATTCAATGCCGTG  
 AGTTTCTTCTCACGCTGCTTGAGCGCGTTCGACCAGCGTAACCCGTAGTTACGGCAC  
  
 AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValVal  
 542 GCCTACTACCGCGGTCTTGACGTGTCGTCATCCCGACCAGCGCGATGTTGTCGTCGTG  
 CGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAGCAGCAC  
 ^ ^  
 550 SAC2, 560 DRD1,  
  
 AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn  
 602 GCAACCGATGCCCTCATGACCGGCTATACCGGCAGTCGACTCGGTGATAGACTGCAAT  
 CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA  
 ^  
 615 BSPH1,  
  
 ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle

## FIGURE 14 - Page

662 ACGTGTGTCACTTACAGACAGTCGATTCAGCCTTGACCTACCTTCACCATGAGACAAATC  
 TGCACACAGTGGGTCTGTCAGCTAAAGTCGAACTGGGATGGAAGTGGTAACCTGTTAG  
 ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys  
 722 ACGCTCCCCCAAGATGCTGTCTCCGCCTCAACGTCGGGGCAGGACTGGCAGGGGGAAAG  
 TCGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCTGACCGTCCCCCTTC  
 ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer  
 782 CCAGGCATCTACAGATTTGTGGCACCGGGGAGCGCCCTCCGGCATGTTGACTCGTCC  
 GGTCCTAGATGTCTAAACACCCTGGCCCCCTCGGGGAGGCCGTACAAGCTGAGCAGG  
 816 BGLI, 833 DRD1,  
 ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr  
 842 GTCTCTGTGAGTGCTATGACGCAGGCTGTGCTGGTATGAGCTCACGCCCGCCGAGACT  
 CAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCAGGCGGCTCTGA  
 881 SACI,  
 ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu  
 902 ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGCTTCCGTGTGCCAGGACCATCTT  
 TGTCAATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCTGGTAGAA  
 931 SMAI XMAI,  
 GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln  
 962 GAATTTGGGAGGGCGTCTTACAGGCCTCACTCATATAGATGCCACTTCTATCCCAG  
 CTTAAACCCCTCCCGCAGAAATGTCCGGAGTGAATATCTACGGGTGAAAGATAAGGGTC  
 985 STUI,  
 ThrLysGinSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla  
 1022 ACAAAAGCAGAGTGGGAGAACCTTCCCTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT  
 TGTTCGTCTACCCCTTTGGAAGGAATGGACCATCGCATGGTGGCACACCGCGA  
 1069 DRA3,  
 ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys  
 1082 AGGGCTCAAGCCCTCCCCCATCGTGGGACCAGATGTGGAAGTGTGTTGATTGCGCTCAAG  
 TCCCGAGTTGGGGAGGGGTAGCACCCCTGGTACACCTCACAAACTAACGGAGTTC  
 1142 ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle  
 CCCACCCCTCCATGGGCCAACACCCCTGCTATACAGACTGGCGCTGTTCAGAATGAAATC  
 GGGTGGGAGGTACCCGGTGTGGGACGATATGTCTGACCCCGACAAGTCTTACTTAG  
 1150 NCOI,  
 ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal  
 1202 ACCCTGACGCACCCAGTCACCAAATACATCATGACATGCGATGTCGGCCGACCTGGAGGTC  
 TGGGACTGCGTGGGTCACTGGTTATGTAGTACTGTACGTACAGCCGCTGGACCTCCAG  
 1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,  
 ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu  
 1262 GTCACGAGCACCTGGGTGCTCGTTGGCGCGTCTGGCTGCTTGGCCGCGTATTGCGCTG  
 CAGTGCTCGTGGACCCACGAGCAACCGCCGAGGACCGACGAAACCGGCGCATAACGGAC

## FIGURE 14 - Page 3

1322 SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle  
 TCAACAGGCTGCGTGGTCATAGTGGCAGGGTCGTCTTGTCCGGGAAGCCGGCAATCATA  
 AGTTGTCCGACGCACCAGTATCACCCGTCCCAGCAGAACAGGCCCTCGGCCGTTAGTAT  
 1369 NAEI,  
 ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu  
 1382 CCTGACAGGGAAAGTCCTCTACCGAGAGTCGATGAGATGGAAGAGTGCTCTCAGCACTTA  
 GGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCACGAGAGTCGTGAAT  
 1385 DRD1,  
 ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu  
 1442 CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAAGGCCCTCGGCCTC  
 GGCATGTAGCTCGTCCCTACTACGAGCGGCTCGTCAAGTCTCGTCTCCGGAGGCCGGAG  
 LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln  
 1502 CTGCAGACCGCGTCCCGTCAGGCAGAGGTTATGCCCTGCTGTCCAGACCAACTGGCAA  
 GACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT  
 1502 PSTI, 1507 TTH3I,  
 LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu  
 1562 AAACTCGAGACCTTCTGGCGAAGCATATGTGAACTTCATCAGTGGATACAATACTTG  
 TTTGAGCTCTGGAAGACCGCTTCGTATACACCTTGAAGTAGTCACCCATGTTATGAAC  
 1565 XHOI, 1586 NDEI,  
 AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla  
 1622 GCGGGCTTGTCAACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTACAGCT  
 CGCCCGAACAGTTGCGACGGACCATTGGGCGGTAACGAAGTAACTACCGAAAATGTCGA  
 1643 BSTE2, 1677 ALWN1 PVU2,  
 AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp  
 1682 GCTGTCACCAGCCCACTAACCACTAGCAAACCCCTCTTCACATATTGGGGGGGTGG  
 CGACAGTGGTCGGTGATTGGTATCGGTTGGAGGAGAAGTTGTATAACCCCCCCCACC  
 ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla  
 1742 GTGGCTGCCAGCTCGCCGCCCCGGTGCCTACTGCCTTGTGGCGCTGGCTTAGCT  
 CACCGACGGGTCGAGCGGGGGGCCACGGCGATGACGGAAACACCCCGGACCGAATCGA  
 1794 ESP1,  
 GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr  
 1802 GGCGCCGCATCGCAGTGGACTGGGAAGGTCTCATAGACATCCTGCAGGGTAT  
 CCGCGCGGTAGCCGTACAACCTGACCCCTCCAGGAGTATCTGTAGGAACGTCCCATA  
 1802 KAS1 NARI,  
 GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer  
 1862 GGCGCCGGCGTGGCGGGAGCTTGTGGCATTCAGATCATGAGCGGTGAGGTCCCTCC  
 CCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTCTAGTACTGCCACTCCAGGGAGG  
 1878 SAC1, 1899 BSPH1,

## FIGURE 14 - Page 4

1922 ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly  
 ACGGAGGACCTGGTCAATCTACTGCCGCCATCCTCTGCCGGAGCCCTCGTAGTCGGC  
 TGCCTCCTGGACCAGTTAGATGACGGCGGTAGGAGAGCGGGCCTCGGGAGCAGCCG  
 ^  
 1928 TTH3I,  
 ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp  
 1982 GTGGTCTGTGCAGCAATACTGCAGCCGGCACGTTGGCCCGGGCAGGGGGCAGTGCAGTGG  
 CACCAAGACACGTCGTTATGACGCCGTGCAACCAGGGCCGCTCCCCGTACAGTCACC  
 ^  
 2004 NAEI, 2017 SMAI XMAI,  
 MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal  
 2042 ATGAACCGGCTGATAGCCTTCGCCTCCGGGGAACATGTTCCCCCACGCACACTACGTG  
 TACTTGGCCGACTATCGGAAGCGGGAGGGCCCCCTGGTACAAAGGGGGTGCCTGATGCAC  
 ^  
 2067 SMAI XMAI, 2093 DRA3,  
 ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln  
 2102 CCGGAGAGCGATGCAGCTGCCGCCTGCACGCCACTCAGCAGCCTCACTGTAACCCAG  
 GGCCTCTCGCTACGTCGACGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGGTC  
 ^  
 2115 PVU2, 2159 ALWN1,  
 LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer  
 2162 CTCCTGAGGCAGACTGCACCAGTGGATAAGCTCGGAGTGTACCACTCCATGCTCCGGTTCC  
 GAGGACTCCGCTGACGTGGTACCTATTGAGCCTCACATGGTGAGGTACGAGGCCAAGG  
 ^  
 2164 MST2, 2220 ECON1,  
 TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu  
 2222 TGGCTAAGGGACATCTGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACCTGGCTA  
 ACCGATTCCCTGTAGACCTGACCTATACGCTCCACAACTCGCTGAAATTCTGGACCGAT  
 ^  
 LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr  
 2292 AAAGCTAAGCTCATGCCACAGCTGCCCTGGATCCCCTTGTGTCCTGCCAGCGCGGGTAT  
 TTTCGATTGAGTACGGTGTGACGGACCCCTAGGGAAACACAGGACGGTCGCGCCCATA  
 ^  
 2285 ESP1, 2300 PVU2, 2310 BAMHI,  
 LysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAlaGluIle  
 2342 AAGGGGGTCTGGCAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC  
 TTCCCCCAGACCGCTCCCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAG  
 ^  
 ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet  
 2402 ACTGGACATGTCAAAACGGGACGATGAGGATCGTCGGTCTAGGACTCTGAGAACATG  
 TGACCTGTACAGTTTGCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCTGTAC  
 ^ ^ ^  
 2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,  
 TrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeuProAla  
 2462 TGGAGTGGGACCTCCCCATTAATGCCTACACCAACGGGCCCCGTACCCCCCTTCCTGCG  
 ACCTCACCCCTGGAAGGGTAATTACGGATGTGGTGCCCCGGGACATGGGGGAAGGACGC  
 ^  
 2480 ASE1, 2497 APAI,  
 ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln

## FIGURE 14 - Page 5

2522 CCGAACTACACGTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATAAGGCAG  
GGCTGATGTGCAAGCGCGATACCTCCCACAGACGTCTCCTTATGCACCTCTATTCCGTC  
2553 PSTI,  
ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln  
2582 GTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCGTGCCAG  
CACCCCCCTGAAGGTGATGCACTGCCCATACTGATGACTGTAGAATTACGGGCACGGTC  
2594 DRA3,  
ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro  
2642 GTCCCATGCCCGAATTTTACAGAAATTGGACGGGTGCCCTACATAGGTTGCCCG  
CAGGGTAGCGGGCTTAAAAAGTGTCTAACCTGCCACCGGGATGTATCCAAACGCCGG  
ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro  
2702 CCCTGCAAGCCCTTGCTGCCGGAGGAGGTATCATTAGAGTAGGACTCCACGAATACCCG  
GGGACGTTGGAAACGACGCCCTCCATAGTAAGTCTCATCCTGAGGTGCTTATGGC  
2757 HGIE2,  
ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu  
2762 GTAGGGTCGCAATTACCTTGCAGCCCGAACCGGACGTGGCGTGTGACGTCCATGCTC  
CATCCCAGCGTTAATGGAACGCTCGGGCTTGGCGTGCACCGGCACAACTGCAGGTACGAG  
2809 AAT2,  
ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro  
2822 ACTGATCCCTCCCATATAACAGCAGAGGCCGGCGAACGGTTGGCGAGGGGATCACCC  
TGACTAGGGAGGGTATATTGCGTCTCCGCCGGCGCTTCCAACCGCTCCCTAGTGGG  
2850 EAG1 XMA3,  
ProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAlaThrCys  
2882 CCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCAAGGCAACTTGC  
GGGAGACACCGGTCGAGGAGCCGATCGGTGATAGGCGAGGTAGAGAGTTCCGTTGAACG  
2889 BALI, 2903 NHEI,  
ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln  
2942 ACCGCTAACCATGACTCCCTGATGCTGAGCTCATAGAGGCCAACCTCTATGGAGGCAG  
TGGCGATTGGTACTGAGGGACTACGACTCGAGTATCTCCGGTGGAGGATACCTCCGTC  
2966 ESP1, 2969 SACI,  
GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer  
3002 GAGATGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTATTCTGGACTCC  
CTCTACCCGCCGTTGATGGTCCCAACTCAGTCTTGTTCACCAACTAACGACCTGAGG  
PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu  
3062 TTGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCTG  
AAGCTAGGCGAACACCGCCTCCTGCTGCCCTCTAGAGGCATGGCGTCTTAGGAC  
3096 BGL2,  
ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro  
3122 CGGAAGTCTCGGAGATTGCCAGGCCCTGCCGTTGGCGCGGGACTATAACCC

## FIGURE 14 - Page 6

GCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGGCCTGATATTGGG  
 3143 ALWN1, 3164 EAG1 XMA3,  
 ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro  
 3182 CCCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACCACCTGTGGTCCATGGCTGCCCG  
 GGCGATCACCTCTGCACCTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCGACGGGC  
 3217 HGIE2, 3229 NCOI,  
 LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu  
 3242 CTTCCACCTCAAAGTCCCTCCTGTGCCCTCGCCTCGGAAGAAGCGGACGGTGGTCCTC  
 GAAGGTGGAGGTTTCAGGGAGGACACGGAGGCGGAGCCTTCTCGCCTGCCACCAGGAG  
 ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer  
 3302 ACTGAATCAACCCTATCTACTGCCTGGCGAGCTCGCCACCAGAAAGCTTGGCAGCTCC  
 TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGCTTCGAAACCGTCGAGG  
 3332 SACI, 3346 HIND3,  
 SerThrSerGlyIleThrGlyAspAsnThrThrSerSerGluProAlaProSerGly  
 3362 TCAACTTCCGGCATTACGGCGACAATAACGACAACATCCTCTGAGCCGCCCTCTGGC  
 AGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTAGGAGACTCGGGCGGGAAAGACCG  
 CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro  
 3422 TGCCCCCCCAGCTCCGACGCTGAGTCTATTCCCTCATGCCCTGGAGGGGGAGCCT  
 ACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGACCTCCCCCTCGGA  
 3437 EAM11051,  
 GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu  
 3482 GGGGATCCGGATCTTAGCGACGGGTATGGTCACGGTCAGTAGTGAGGCCAACGGAG  
 CCCCTAGGCCTAGAATGCTGCCAGTACAGTGCAGTCACACTCCGGTTGCGCCTC  
 3484 BAMHI, 3485 BSAB1, 3487 BSPE1,  
 AspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla  
 3542 GATGTCGTGCTGCTCAATGTCTTACTCTTGGACAGGCAGCTCGTCACCCGTGCGCC  
 CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCGGCGTGAGCAGTGGGGCACGCG  
 3589 DRA3, 3600 SAC2,  
 AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn  
 3602 GCGGAAGAACAGAAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTACCCACAAT  
 CGCCTTCTTGTCTTGACGGGTAGTTACGTGATTGAGCAACGATGCAGTGGTGT  
 3611 ALWN1, 3655 PFLM1,  
 LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp  
 3662 TTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAGAAAGTCACATTGAC  
 AACCACATAAGGTGGAGTGCCTCACGAACGGTTCCGTCTTCAGTGTAAACTG  
 3681 DRA3,  
 ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla  
 3722 AGACTGCAAGTTCTGGACAGCCATTACCAAGGACGTACTCAAGGAGGTAAAGCAGCGCG  
 TCTGACGTTCAAGACCTGTCGGTAATGGTCTGCATGAGTCCCTCCAATTTCGTCGCCGC

## FIGURE 14 - Page 7

SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis  
 3782 TCAAAAGTGAAGGCTAACCTGCTATCCGTAGAGGAAGCTTGCAGCCTGACGCCACAC  
 AGTTTCACTTCCGATTGAACGATAGGCATCTCCTCGAACGTCGGACTGCGGGGTGTG  
 ^  
 3816 HIND3,

SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla  
 3842 TCAGCCAAATCCAAGTTGGTATGGGGCAAAAGACGTCCGTTGCCATGCCAGAAAGGCC  
 AGTCGGTTAGGTTCAAACCAATACCCCGTTCTGCAGGCAACGGTACGGTCTTCCGG  
 ^  
 3875 AAT2, 3890 BGLI,

ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp  
 3902 GTAAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCAATAGAC  
 CATTGGGTGTAGTTGAGGCACACCTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG  
 ^  
 3962 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyArgLys  
 ACTACCACATGGCTAAAGAACGAGGTTCTGCGTCAGCCTGAGAAGGGGGTCGTAAG  
 TGATGGTAGTACCGATTCTGCTCCAAAGACGCAAGTCGGACTCTCCCCCAGCATTG  
 ^  
 4022 ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu  
 CCAGCTCGTCTCATCGTGTTCGGCGATCTGGCGTGCCTGCGAAAGATGGCTTGG  
 GGTGAGCAGAGTAGCACAAAGGGCTAGACCCGCACCGCACACGCTTCTACCGAAAC  
 ^  
 4082 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr  
 TACGACGTGGTTACAAAGCTCCCTGGCCGTGATGGGAAGCTCCTACGGATTCCAATAC  
 ATGCTGCACCAATGTTGAGGGGAAACGGCACTACCCCTCGAGGATGCCCTAAGGTTATG  
 ^  
 4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet  
 TCACCAGGACAGCGGGTTGAATTCTCGTCAAGCGTGGAAAGTCCAAGAAAACCCAATG  
 AGTGGTCTGTCGCCAACTTAAGGAGCACGTTCGCACCTCAGGTTCTGGGTAC  
 ^  
 4160 ECORI,

GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr  
 4202 GGGTCTCGTATGATACCGCTGCTTGACTCCACAGTCACTGAGAGCGACATCCGTACG  
 CCCAAGAGCATACTATGGCGACGAAACTGAGGTGTCAGTGACTCTCGCTGTAGGCATGC  
 ^  
 4229 DRD1, 4236 ALWN1,

GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer  
 4262 GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCCCAAGCCCGTGGCCATCAAGTCC  
 CTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGTTCGGCGCACCGTAGTCAGG  
 ^  
 4301 BGLI, 4308 BALI,

LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly  
 4322 CTCACCGAGAGGCTTATGTTGGGGCCCTCTTACCAATTCAAGGGGGAGAACCTGCGGC  
 GAGTGGCTCTCCGAAATACAACCCCGGGAGAATGGTTAACACTGGAGCTGGGGTTCAGG  
 ^  
 4345 APAI,

TyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCys  
 4382 TATCGCAGGTGCCGCGCGAGCGGGTACTGACAACTAGCTGTGGTAACACCCTACTTGC  
 ATAGCGTCCACGGCGCGCTCGCCGATGACTGTTGATCGACACCATTGTGGAGTGAACG

## FIGURE 14 - Page 8

TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal  
 4442 TACATCAAGGCCGGGCAGCCTGTCGAGCCGAGGGCTCCAGGACTGCACCATGCTCGTG  
 ATGTAGTTCCGGGCCGTCGGACAGCTCGCGTCCCAGGGTCTGACGTGGTACGAGCAC  
 ^  
 4452 SMA1 XMA1,  
  
 CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer  
 4502 TGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGTCCAGGAGGACGCAGCGAGC  
 ACACCGCTGCTGAATCAGCAATAGACACTTCGCGCCCCCAGGTCTCCTGCGCCGCTCG  
 ^ ^  
 4508 DRD1, 4511 TTH3I,  
  
 LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln  
 4562 CTGAGAGCCTTCACGGAGGCTATGACCAAGGTACTCCGCCCCCTGGGGACCCCCCACAA  
 GACTCTCGGAAGTGCCTCGATACTGGTCCATGAGGCGGGGGGACCCCTGGGGGGTGT  
  
 ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp  
 4622 CCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCAAACGTGTCAGTCGCCACGAC  
 GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGGGTGT  
 ^  
 4637 SAC1,  
  
 GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla  
 4682 GGCCTGGAAAGAGGGTCTACTACCTCACCGTGACCCCTACAACCCCCCTCGCGAGAGCT  
 CCGCGACCTTCTCCAGATGATGGAGTGGGCACTGGATGTTGGGGAGCGCTCTCGA  
 ^  
 4731 NRUI,  
  
 AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe  
 4742 GCGTGGGAGACAGCAAGACACACTCCAGTCATTCTGGCTAGGCAACATAATCATGTTT  
 CGCACCCCTCTGCTGTTCTGTGAGGTCAAGGACCGATCCGTTGTATTACTACAAA  
  
 AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla  
 4802 GCCCCCACACTGTGGCGAGGATGATACTGATGACCCATTCTTTAGCGCCTTATAGCC  
 CGGGGGTGTGACACCCGCTCCTACTATGACTACTGGTAAAGAAATCGCAGGAATATCGG  
 ^ ^  
 4806 PFLM1, 4807 DRA3,  
  
 ArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu  
 4862 AGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGCCTGCTACTCCATAGAA  
 TCCCTGGTCGAACCTGTCCGGAGCTAACGCTCTAGATGCCCGGACGATGAGGTATCTT  
 ^  
 4893 BGL2,  
  
 ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis  
 4922 CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTCACTCCAC  
 GGTGACCTAGATGGAGGTTAGTAAGTTCTGAGGTACCGGAGTCGCGTAAAGTGAGGTG  
 ^  
 4954 NCOI,  
  
 SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro  
 4982 AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAAACCTGGGGTACCG  
 TCAATGAGAGGTCCACTTAGTTATCCCACCGGGTACGGAGTCTTTGAACCCCATGGC  
 ^ ^  
 5015 SPHI, 5035 KPNI,  
  
 ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly

## FIGURE 14 - Page 9

5042 CCCTTGCAGCTGGAGACACCAGGGCCGGAGCGTCCGCGTAGGCTTCTGGCCAGAGGA  
GGAACGCTCGAACCTCTGTGGCCGGGCTCGCAGGCGATCCGAAGACCGTCTCCT  
5064 APAI, 5091 BALI,  
GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys  
5102 GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGCAGTAAGAACAAAGCTCAAA  
CCGTCGGACGGTATACACCGTTATGGAGAAGTTGACCCGTATTCTGTTGAGTT  
5113 NDEI,  
LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr  
5162 CTCACTCCAATAGCGGCCCTGGCCAGCTGGACTTGTCCGGCTGGTCACGGCTGGCTAC  
GAGTGAGGTTATGCCGGCGACCGGTCGACCTGAACAGGCCGACAAAGTGCCGACCGATG  
5174 NOTI, 5175 EAG1 XMA3, 5182 BALI, 5186 PVU2,  
SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys  
5222 AGCGGGGGAGACATTATCACAGCGTGTCTCATGCCGGCCCGCTGGATCTGGTTTG  
TCGCCCCCTCTGTAAATAGTGTGCGCACAGAGTACGGGCCGGCGACCTAGACCAAAACG  
5240 DRA3,  
LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgOP  
5282 CTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCCAACCGATGAATAGTCGAC  
GATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGTTGGCTACTTATCAGCTG  
5295 PSTI, 5336 SALI,

FIGURE 15

SCANNED, # 14



ROBINS & ASSOCIATES  
90 Middlefield Road, Ste 200  
Menlo Park, CA 94025  
USSN 2302-1617  
Sheet 54 of 100

FIGURE 16 - Page 1

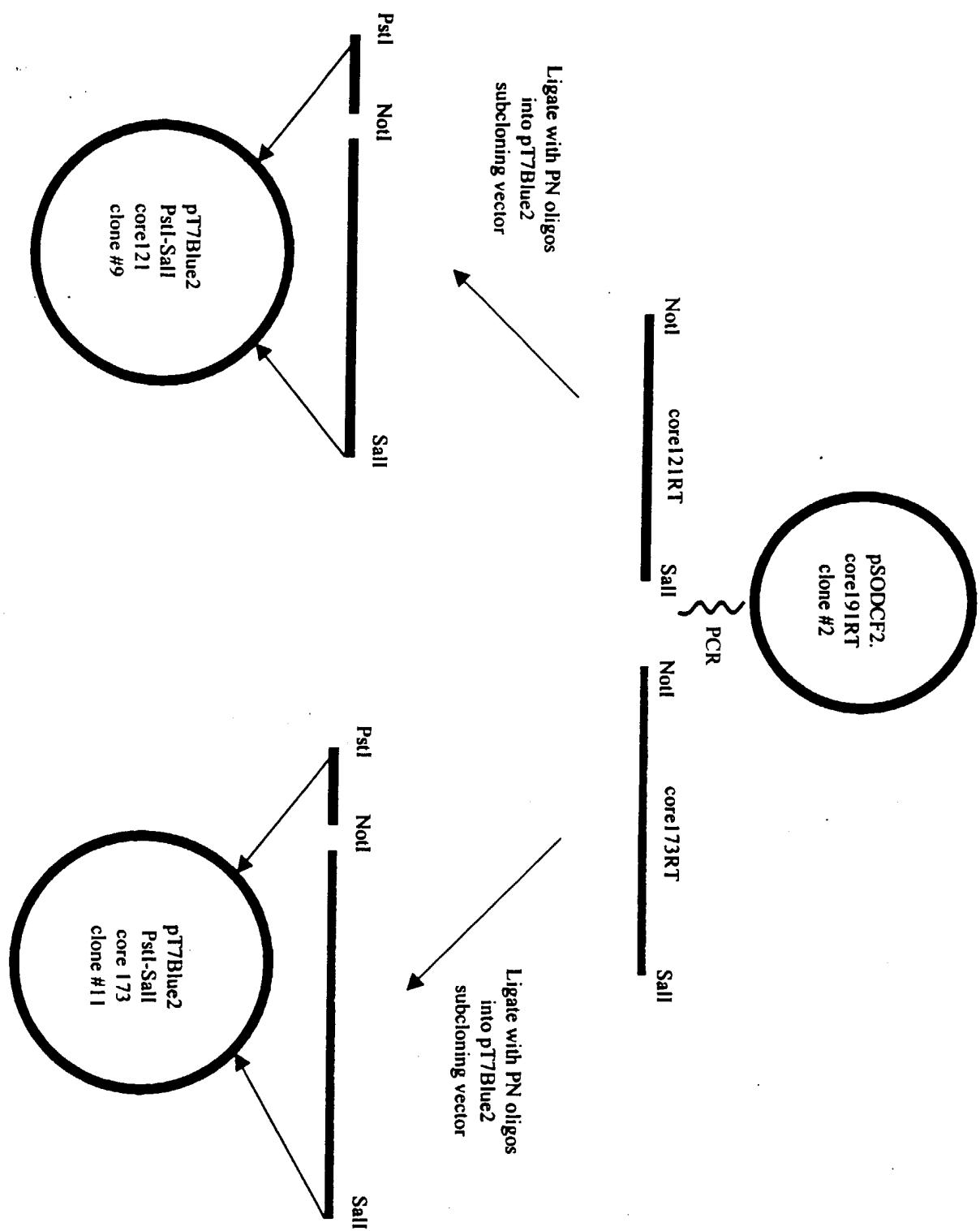
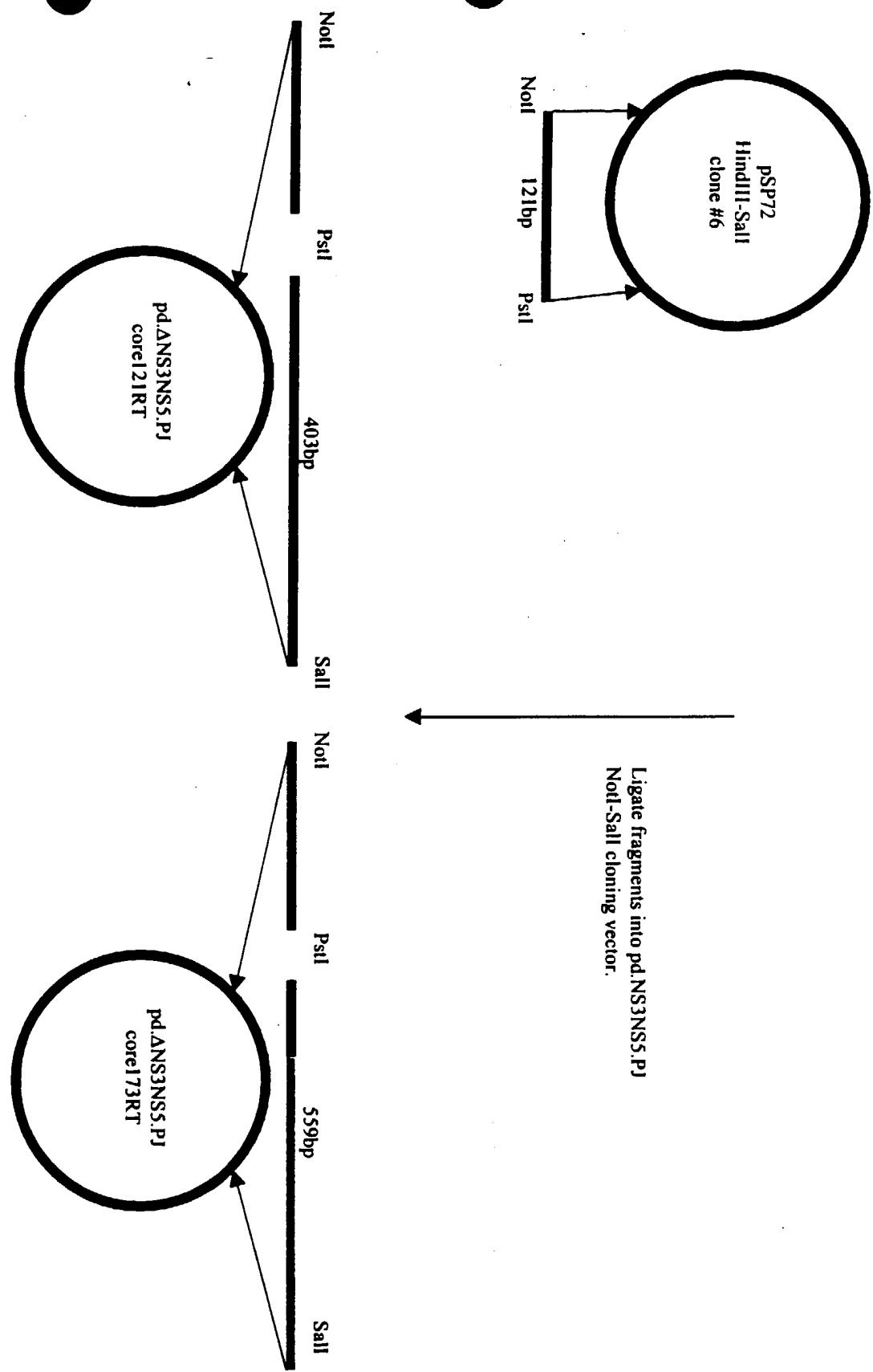


FIGURE 16 - Pa



Ligate fragments into pd.NS3NS5.PJ  
NotI-Sall cloning vector.

## FIGURE 17 - Page 1

Met Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn  
2 AGCTTACAAAACAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC  
TCGAATGTTGTTTACCGACGTATACGTCGAGTCCCATAATTCCACGATCATGAGTTG  
^ ^ ^  
1 HIND3, 24 NDEI, 52 SCAI,

Pro Ser Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr Met Ser Lys Ala His Gly Ile Asp  
62 CCCTCTGTTGCTGCAACACTGGGCTTGGTCTTACATGTCCAAGGCTCATGGGATCGAT  
GGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCCTAGCTA  
^  
116 CLAI,

Pro Asn Ile Arg Thr Gly Val Arg Thr Ile Thr Thr Gly Ser Pro Ile Thr Tyr Ser Thr  
122 CCTAACATCAGGACCGGGGTGAGAACAAATTACCACTGGCAGCCCCATACGTACTCCACC  
GGATTGTAGTCCTGGCCCCACTCTTGTAAATGGTGACCGTCGGGTAGTGCATGAGGTGG

Tyr Gly Lys Phe Leu Ala Asp Gly Gly Cys Ser Gly Ala Tyr Asp Ile Ile Ile Cys  
182 TACGGCAAGTTCTGCCGACGGCGGGTGCTCGGGGGCGCTTATGACATAATAATTGT  
ATGCCGTTCAAGGAACGGCTGCCGCCACGAGCCCCCGCGAATACTGTATTATAAACAA

Asp Glu Cys His Ser Thr Asp Ala Thr Ser Ile Leu Gly Ile Gly Thr Val Leu Asp Gln  
242 GACGAGTGCCACTCCACGGATGCCACATCCATCTGGCATTGGCACTGTCCTTGACCAA  
CTGCTCACGGTGAGGTGCTACGGTGTAGGTAGAACCGTAACCGTACAGGAACGGTT

Ala Glu Thr Ala Gly Ala Arg Leu Val Val Leu Ala Thr Ala Thr Pro Pro Gly Ser Val  
302 GCAGAGACTGCAGGGGGCGAGACTGGTTGTGCTGCCACGCCACCCCTCCGGGCTCGTC  
CGTCTCTGACGCCCGCTCTGACCAACACGAGCGGTGGCGTGGGGAGGCCCGAGGCAG  
^  
303 ALWN1,

Thr Val Pro His Pro Asn Ile Glu Glu Val Ala Leu Ser Thr Thr Gly Glu Ile Pro Phe  
362 ACTGTGCCCATCCAACATCGAGGAGGTTGCTCTGTCCACCACCGGAGAGATCCCTTT  
TGACACGGGGTAGGGTTAGCTCCTCAACGAGACAGGTGGTGGCCTCTAGGGAAAA

Tyr Gly Lys Ala Ile Pro Leu Glu Val Ile Lys Gly Gly Arg His Leu Ile Phe Cys His  
422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGAGACATCTCATCTTGTGTCAT  
ATGCCGTTCCGATAGGGGGAGCTTCATTAGTCCCCCCCCTGTAGAGTAGAACAGTA

## FIGURE 17 - Page 2

482 SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal  
 TCAAAAGAAGAGTCGACGAACTCGCCGCAAAGCTGGTCGATTGGGCATCAATGCCGTG  
 AGTTTCTTCTCACGCTGCTTGAGCGCGTTCGACCAGCGTAACCCGTAGTTACGGCAC

542 AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal  
 GCCTACTACCGCGGTCTTGACGTGTCGTACCCGACCAGCGCGATGTTGTCGTCGTG  
 CGGATGATGGCGCCAGAACCTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAGCAGCAC  
 ^ ^  
 550 SAC2, 560 DRD1,

602 AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn  
 GCAACCGATGCCCTCATGACCGGCTATAACCGGCACTTCGACTCGGTGATAGACTGCAAT  
 CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA  
 ^  
 615 BSPH1,

662 ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle  
 ACGTGTGTCACCCAGACAGTCGATTCAGCCTTGACCCCTACCTCACCATTGAGACAATC  
 TGCACACAGTGGGTCTGTCAGCTAAAGTCGGAACTGGGATGGAAGTGGTAACTCTGTTAG

722 ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys  
 ACGCTCCCCAAGATGCTCTCCGCCTCAACGTCGGGGCAGGACTGGCAGGGGGAAAG  
 TGCAGGGGGTCTACGACAGAGGGCGTGAGTGCAGCCCCGTCCTGACCGTCCCCCTTC

782 ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer  
 CCAGGCATCTACAGATTGTGGCACCGGGGAGCGCCCTCCGGCATGTTGACTCGTCC  
 GGTCCGTAGATGTCTAACACCGTGGCCCCCTCGGGGGAGGCCGTACAAGCTGAGCAGG  
 ^  
 816 BGLI, 833 DRD1,

842 ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr  
 GTCCTCTGTGAGTGTCTATGACGCAGGCTGTGCTGGTATGAGCTCACGCCGCCGAGACT  
 CAGGAGACACTACGATACTCGTCCGACACGAACCATACTCGAGTGCAGGGCGGCTCTGA  
 ^  
 881 SAC1,

902 ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu  
 ACAGTTAGGCTACGAGCGTACATGAACACCCGGGCTTCCGTGTGCCAGGACCATCTT  
 TGTCAAATCCGATGCTCGCATGTACTTGTGGGGCCCGAAGGGCACACGGCCTGGTAGAA  
 ^  
 931 SMA1 XMA1,

962 GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln  
 GAATTTGGGAGGGCGTCTTACAGGCCTCACTCATATAGATGCCACTTCTATCCCAG  
 CTTAAACCCCTCCGCAGAAATGTCGGAGTGAGTATCTACGGGTGAAAGATAGGGTC  
 ^  
 985 STUI,

1022 ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla  
 ACAAAAGCAGAGTGGGAGAACCTCCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT  
 TGTTCGTCTACCCCTCTGGAAAGGAATGGACCATCGCATGGTCGGTGGCACACGCGA  
 ^  
 1069 DRA3,

1082 ArgAlaGlnAlaProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys  
 AGGGCTCAAGCCCCCTCCCCCATCGTGGGACCAGATGTGAAAGTGGTTGATTGCCCTCAAG

# FIGURE 17 - Page 3

TCCCGAGTTGGGGAGGGGGTAGCACCCCTGGTCTACACCTCACAAACTAAGCGGAGTTC  
 1142 ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle  
 CCCACCCCTCCATGGGCCAACACCCCTGCTATACAGACTGGCGCTGTTAGAATGAAATC  
 GGGTGGGAGGTACCCGGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTAG  
 ^  
 1150 NCOI,  
 1202 ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal  
 ACCCTGACGCACCCAGTCACCAATACATCATGACATGCGATGTCGGCCGACCTGGAGGTC  
 TGGGACTGCGTGGGTCACTGGTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG  
 ^ ^ ^ ^ ^  
 1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,  
 1262 ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu  
 GTCACAGCAGCCTGGGTGCTCGTGGCGCGCTGGCTGCTTGGCCGCGTATTGCCCTG  
 CAGTGCTCGTGGACCCACGAGCAACCGCCGAGGACCGACGAAACCGCCGATAACGGAC  
 1322 SerThrGlyCysValValIleValGlyArgValLeuSerGlyLysProAlaIleIle  
 TCAACAGGCTGCGTGGTCATAGTGGCAGGGTCGTCTTGTCCGGAAAGCCGGCAATCATA  
 AGTTGTCCGACGCACCAGTATCACCCGTCCCAGCAGAACAGGCCCTCGGCCGTAGTAT  
 ^  
 1369 NAEI,  
 1382 ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu  
 CCTGACAGGGAAAGTCCTCTACCGAGAGTTGATGAGATGGAAGAGTGCTCTCAGCACTTA  
 GGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCACGAGAGTCGTGAAT  
 ^  
 1385 DRD1,  
 1442 ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu  
 CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAAGGCCCTCGGCCTC  
 GGCATGTAGCTCGTCCCTACTACGAGCGGCTCGTCAAGTTCTCCGGAGGCCGGAG  
 1502 LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln  
 CTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCTGCTGTCCAGACCAACTGGCAA  
 GACGTCTGGCGCAGGGCAGTCCGTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT  
 ^ ^  
 1502 PSTI, 1507 TTH3I,  
 1562 LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu  
 . . . AAACCTCGAGACCTCTGGCGAAGCATATGTGGAACCTTCATCAGTGGGATAACAATACTG  
 TTTGAGCTCTGGAAGACCCGTTCGTATACACCTTGAAGTAGTCACCCCTATGTTATGAAC  
 ^ ^  
 1565 XHOI, 1586 NDEI,  
 1622 AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla  
 GCGGGCTTGTCAACGCTGCCGGTAACCCCGCCATTGCTTCATTGATGGCTTACAGCT  
 CGCCCGAACAGTTGCGACGGACCATTGGGGCGGTAAACGAAGTAACCGAAAATGTCGA  
 ^ ^  
 1643 BSTE2, 1677 ALWN1 PVU2,  
 1682 AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp  
 GCTGTCACCAGCCCCTAACCAACTAGCCAAACCCCTCTCAACATATTGGGGGGGTGG  
 CGACAGTGGTGGGTGATTGGTATCGGTTGGGAGGAGTTGTATAACCCCCCACC

# FIGURE 17 - Page

1742 ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla  
 GTGGCTGCCAGCTGCCGCCCCGGTGCCGCTACTGCCCTTGTGGCGCTGGCTTAGCT  
 CACCGACGGGTCAGCGGGGGCACGGCGATGACGGAAACACCCGCGACCGAATCGA  
  
 1794 ESP1,  
  
 1802 GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr  
 GGCGCCGCATCGGCAGTGGACTGGGAAAGGTCTCATAGACATCCTGCAGGGTAT  
 CCGCGGCGGTAGCGTCACAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCCATA  
 ^  
 1802 KAS1 NARI,  
  
 1862 GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer  
 GGC CGGGCGTGGCGGGAGCTTGTGGCATTCAAGATCATGAGCGGTGAGGTCCCCTCC  
 CCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG  
 ^  
 1878 SAC1, 1899 BSPH1,  
  
 1922 ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly  
 ACGGAGGACCTGGTCAATCTACTGCCCGCCATCCTCTGCCCGAGGCCCTCGTAGTCGGC  
 TGCCTCCTGGACCAGTTAGATGACGGCGGTAGGAGAGCGGGCTCGGAGCATCAGCCG  
 ^  
 1928 TTH3I,  
  
 1982 ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp  
 GTGGTCTGTGCAGCAATACTGCAGCCGGCACGTTGGCCCGGGCAGGGGGCAGTGCAGTGG  
 CACCAGACACGTCGTTATGACGCGGGCGTGCACCGGGCCGCTCCCCCGTCACGTCA  
 ^  
 2004 NAEI, 2017 SMAI XMAI,  
  
 2042 MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal  
 ATGAACCGGCTGATAGCCTTCGCCTCCGGGAACCATGTTCCCCCACGCACACTACGTG  
 TACTTGGCGACTATCGGAAGCGGAGGGCCCCCTGGTACAAAGGGGGTGCCTGATGCAC  
 ^  
 2067 SMAI XMAI, 2093 DRA3,  
  
 2102 ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGin  
 CCGGAGAGCGATGCGAGCTGCCCGCGTCAGTGCCTACTGCCATACTCAGCAGCCTCA  
 GGCCTCTCGCTACGTCGACGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGTC  
 ^  
 2115 PVU2, 2159 ALWN1,  
  
 2162 LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer  
 CTCCTGAGGCAGTGCACCAAGTGGATAAGCTGGAGTGTACCACTCCATGCTCCGGTTCC  
 GAGGACTCCGCTGACGTGGTCACCTATTGAGCCTCACATGGTGAGGTACGAGGCCAAGG  
 ^  
 2164 MST2, 2220 ECON1,  
  
 2222 TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu  
 TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTAAGACCTGGCTA  
 ACCGATTCCCTGTAGACCTGACCTATACGCTCCACAACCTCGCTGAAATTCTGGACCGAT  
  
 2282 LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr  
 AAAGCTAACGCTCATGCCACAGCTGCCCTGGATCCCTTGTGTCTGCCAGCGCGGGTAT  
 TTTGATTGAGTACGGTGTGACGGACCCCTAGGGAAACACAGGACGGTCGCGCCCCATA  
 ^ ^ ^  
 2285 ESP1, 2300 PVU2, 2310 BAMHI,

## FIGURE 17 - Page 5

2342 LysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAlaGluIle  
 AAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC  
 TTCCCCCAGACCGCTCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAG

2402 ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet  
 ACTGGACATGTCAAAACGGGACGATGAGGATCGTCGGCTCTAGGACCTGCAGGAACATG  
 TGACCTGTACAGTTTGCCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCCTGTAC

2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,

2462 TrpSerGlyThrPheProIleAsnAlaTyrThrGlyProCysThrProLeuProAla  
 TGGAGTGGGACCTCCCCATTAATGCCTACACCACGGGCCCCGTACCCCCCTTCCTGCG  
 ACCTCACCCCTGGAAGGGTAATTACGGATGTGGTGCCCCGGGACATGGGGGAAGGACGC

2480 ASE1, 2497 APAI,

2522 ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln  
 CCGAACTACACGTTCGCGCTATGGAGGGTGTGCAGAGGAATACGTGGAGATAAGGCAG  
 GGCTTGATGTGCAAGCGCGATACCTCCCACAGACGTCTCCTTATGCACCTCTATTCCGTC

2553 PSTI,

2582 ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln  
 GTGGGGACTTCACTACGTGACGGGTATGACTACTGACAATCTAAATGCCGTGCCAG  
 CACCCCTGAAGGTGATGCACTGCCATACTGATGACTGTTAGAATTACGGGACGGTC

2594 DRA3,

2642 ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro  
 GTCCCATCGCCCGAATTTTACAGAATTGGACGGGGTGCCTACATAGGTTGCGCCC  
 CAGGGTAGCGGGCTTAAAAAGTGTCTAACCTGCCAACGCGGATGTATCCAAACGCGGG

2702 ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro  
 CCCTGCAAGCCCTTGCTCGGGAGGGAGGTATCATTAGAGTAGGACTCCACGAATACCCG  
 GGGACGTTGGAACGACGCCCTCCATAGTAAGTCTCATCCTGAGGTGCTTATGGGC

2757 HGIE2,

2762 ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu  
 GTAGGGTCGCAATTACCTTGCAGGCCGAACCGGACGTGGCGTGTGACGTCCATGCTC  
 CATCCCAGCGTTAATGGAACGCTCGGGTTGGCCTGCACCGGACAACACTGCAGGTACGAG

2809 AAT2,

2822 ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro  
 ACTGATCCCTCCCATAAACAGCAGAGGGCGGCCGGCGAACGGTTGGCGAGGGGATCACCC  
 TGACTAGGGAGGGTATATTGTCGTCTCCGCCGGCCGCTTCCAACCGCTCCCTAGTGGG

2850 EAG1 XMA3,

2882 ProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAlaThrCys  
 CCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCAAGGCAACTTGC  
 GGGAGACACCGGTCGAGGAGCCGATCGGTGATAGGCAGGTAGAGAGTTCCGTTGAACG

2889 BALI, 2903 NHEI,

# FIGURE 17 - Page 6

2942 ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln  
 ACCGCTAACCATGACTCCCCTGATGCTGAGCTCATAGAGGCCAACCTCTATGGAGGCAG  
 TGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTGGAGGATACTCCGTC  
 ^ ^  
 2966 ESP1, 2969 SACI,  
 3002 GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer  
 GAGATGGCGGCAACATCACCAAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTGGACTCC  
 CTCTACCCGCCGTTGTAGTGGTCCAACTCAGTCTTGTGTTACCACTAAGACCTGAGG  
 3062 PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu  
 TTGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCGCAGAAATCCTG  
 AAGCTAGGCGAACACCGCCTCCTGCTCGCCCTAGAGGCATGGCGTCTTAGGAC  
 ^  
 3096 BGL2,  
 3122 ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro  
 CGGAAAGTCTCGGAGATTGCCAGGCCCTGCCGTTGGCGCGCCGGACTATAACCC  
 GCCTTCAGAGCCTCTAACGGGTCCGGGACGGCAAACCCGCGCCGGCTGATATTGGGG  
 ^ ^  
 3143 ALWN1, 3164 EAG1 XMA3,  
 3182 ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro  
 CCGCTAGTGGAGACGTGGAAAAAGCCGACTACGAACCACTGTGGTCCATGGCTGCCG  
 GGCATCACCTCTGCACCTTTCTGGCTGATGCTGGTGGACACCAGGTACCGACGGC  
 ^ ^  
 3217 HGIE2, 3229 NCOI,  
 3242 LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu  
 CTTCCACCTCCAAAGTCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTGGCCTC  
 GAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTCGCCTGCCACCAGGAG  
 3302 ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer  
 ACTGAATCAACCCTATCTACTGCCTGGCGAGCTGCCACCAAGCTTGGCAGCTCC  
 TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTCGAAACCGTCGAGG  
 ^ ^  
 3332 SACI, 3346 HIND3,  
 3362 SerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly  
 TCAACTTCCGGCATTACGGCGACAATACGACAACATCCTCTGAGCCGCCCTCTGGC  
 AGTTGAAGGCCGTAATGCCCGCTGTATGCTGTTAGGAGACTCGGGGGGAAGACCG  
 3422 CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGluPro  
 TGCCCCCCCAGCTCCGACGCTGAGTCTATTCCCATGCCCGGCTGGAGGGGGAGCCT  
 ACGGGGGGCTGAGGCTCGACTCAGGATAAGGAGGTACGGGGGGACCTCCCCCTCGGA  
 ^  
 3437 EAM11051,  
 3482 GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu  
 GGGGATCCGGATCTTAGCGACGGGTATGGTCAACGGTCAGTAGTGAGGCCAACGCGGAG  
 CCCCTAGGCCTAGAACATCGCTGCCAGTACCGTTGCCAGTCATCACTCCGGTTCGCC  
 ^ ^ ^  
 3484 BAMHI, 3485 BSAB1, 3487 BSPE1,  
 3542 AspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla  
 GATGTCGTGTGCTGCTCAATGTCTTACTCTTGACAGGCCACTCGTCACCCCGTGC  
 CTACAGCACACGACGAGTTACAGAACCTGTCCCGTGTGAGCAGTGGGGCACGCGG

## 8

### FIGURE 17 - Page 7

3589 DRA3, 3600 SAC2,

3602 AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn  
GCGGAAAGAACAGAAAATGCCCATCAATGCACTAAGCAACTCGTTGCTACGTACCCACAAT  
CGCCTTCTTGTCTTGACGGTAGTTACGTGAGCAACGATGCAGTGGTGTAA

3611 ALWN1, 3655 PFLM1,

3662 LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp  
TTGGTGTATTCCACCACCTCACGCAGTGCTTGCAAAGGCAGAAGAACATTTGAC  
AACCACATAAGGTGGTAGTGCACGAACGGTTCCGTCTTCAGTGTAAACTG

3681 DRA3,

3722 ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla  
AGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCAGCGCG  
TCTGACGTTCAAGACCTGTCGGAATGGCCTGCATGAGTCCCTCCAATTCGTGCCGC

3782 SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis  
TCAAAAGTGAAGGCTAACTGCTATCCGTAGAGGAAGCTTGCAGCCTGACGCCACAC  
AGTTTCACTCCGATTGAACGATAGGCATCTCCTCGAACGTCGGACTGCGGGGGTGTG

3816 HIND3,

3842 SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla  
TCAGCCAAATCCAAGTTGGTTATGGGGCAAAAGACGTCCGTTGCCATGCCAGAAAGGCC  
AGTCGTTAGGTTCAAACCAATACCCGTTCTGCAGGCAACGGTACGGTCTTCCGG

3875 AAT2, 3890 BGLI,

3902 ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp  
GTAACCCACATCAACTCCGTGTGGAAAGACCTCTGGAAGACAATGTAACACCAATAGAC  
CATTGGGTGTAGTTGAGGCACACCTTCTGGAAGACCTCTGTACATTGTGGTTATCTG

3962 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyArgLys  
ACTACCACATGGCTAAGAACGAGGTTTCTGCCTGAGAAGGGGGTCTGTAAG  
TGATGGTAGTACCGATTCTGCTCCAAAGACGCAAGTCGGACTCTCCCCCAGCATTC

4022 ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu  
CCAGCTCGTCTCATCGTGTCCCCGATCTGGCGTGCCTGCGAAAAGATGGCTTG  
GGTCGAGCAGAGTAGCACAAAGGGCTAGACCCGCACGCGCACACGCTTCTACCGAAAC

4082 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr  
TACGACGTGGTTACAAAGCTCCCTGGCGTGTGGAAAGCTCCTACGGATTCCAATAC  
ATGCTCACCAATGTTCGAGGGAACCGGACTACCCCTCGAGGATGCCTAACGGTTATG

4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet  
TCACCAAGGACAGCGGGTTGAATTCTCGTGCAGCGTGGAAAGTCCAAGAAAACCCAATG  
AGTGGCCTGTCGCCAACTTAAGGAGCACGTTGCACCTTCAGGTTCTTGGGTTAC

4160 ECORI,

4202 GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr  
GGGTTCTCGTATGATACCCGCTGTTGACTCCACAGTCAGTGAGAGCGACATCCGTACG  
CCCAAGAGCATACTATGGCGACGAAACTGAGGTGTCAGTGACTCTCGCTGTAGGATGC

## FIGURE 17 - Page 8

4229 DRD1, 4236 ALWN1,

4262 GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer  
GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCCAAGCCCGCGTGGCCATCAAGTCC  
CTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTCGGCGCACCGGTAGTTAGG

4301 BGLI, 4308 BALI,

4322 LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly  
CTCACCGAGAGGGTTTATGTTGGGGGCCCTTACCAATTCAAGGGGGAGAACTGCGGC  
GAGTGGCTCTCCGAAATACAACCCCCGGGAGAATGGTTAAGTCCCCCTTTGACGCCG

4345 APAI,

4382 TyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCys  
TATCGCAGGTGCCGCGAGCGCGTACTGACAACTAGCTGTGGTAACACCCCTCACTTGC  
ATAGCGTCCACGGCGCGCTCGCCGATGACTGTTGATCGACACCATTGTGGGAGTGAACG

4442 TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal  
TACATCAAGGCCGGGCAGCCTGTCGAGCCGCAGGGCTCCAGGACTGCACCATGCTCGTG  
ATGTAGTCCGGGCCGTCGGACAGCTGGCGTCCCGAGGTCTGACGTGGTACGAGCAC

4452 SMAI XMAI,

4502 CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer  
TGTGGCGACGACTTAGCTTATCTGTGAAAGCGCGGGGTCAGGAGGACGCGGCGAGC  
ACACCGCTGCTGAATCAGCAATAGACACTTCGCGCCCCAGGTCTCCTGCGCCGCTCG

4508 DRD1, 4511 TTH3I,

4562 LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln  
CTGAGAGCCTTCACGGAGGCTATGACCGAGGTACTCCGCCCCCTGGGACCCCCCACAA  
GACTCTCGGAAGTGCCTCCGATACTGGTCATGAGGCGGGGGGACCCCTGGGGGTGTT

4622 ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp  
CCAGAATACGACTTGGAGGCTATAACATCATGCTCCTCAACGTGTAGTCGCCACGAC  
GGTCTTATGCTGAACCTCGAGTATTGAGTACGAGGAGGTTGCACAGTCAGCGGGTGCTG

4637 SACI,

4682 GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla  
GGCGCTGGAAAGAGGGTCTACTACCTCACCGTGACCCCTACAACCCCCCTCGCGAGAGCT  
CCCGCACCTTCTCCAGATGATGGAGTGGGACTGGGATGTTGGGGAGCGCTCTCGA

4731 NRUI,

4742 AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe  
GCGTGGGAGACAGCAAGACACACTCCAGTCATTCTGGCTAGGCAACATAATCATGTT  
CGCACCCCTCTGCGTTCTGTGAGGTCAAGTAAAGGACCGATCCGTTGTTATTAGTACAAA

4802 AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla  
GCCCCACACTGTGGCGAGGATGATACTGATGACCCATTCTTAGCGTCCTTATAGCC  
CGGGGGTGTGACACCCGCTCCTACTATGACTACTGGTAAAGAAATCGCAGGAATATCGG

4806 PFLM1, 4807 DRA3,

ArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu

## FIGURE 17 - Page 9

4862 AGGGACCAGCTGAACAGGCCCTCGATTGCGAGATCTACGGGGCTGCTACTCCATAGAA  
 TCCCTGGTCGAACTTGTCCGGAGCTAACGCTCTAGATGCCCGGACGATGAGGTATCTT  
 ^  
 4893 BGL2,  
 ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis  
 4922 CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTCACTCCAC  
 GGTGACCTAGATGGAGGTTAGTAAGTTCTGAGGTACCGGAGTCGCGTAAAGTGAGGTG  
 ^  
 4954 NCOI,  
 SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro  
 4982 AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAAACCTGGGGTACCG  
 TCAATGAGAGGTCCACTTAGTTATCCCACCGGCGTACGGAGTCTTGAAACCCATGGC  
 ^  
 5015 SPHI, 5035 KPNI,  
 ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly  
 5042 CCCTTGCAGCTTGGAGACACCGGGCCGGAGCGTCCGCCTAGGCTCTGGCCAGAGGA  
 GGGAACGCTCGAACCTCTGTGGCCGGCGCTCGCAGGCGCAGTCCGAAGACCGGTCTCCT  
 ^  
 5064 APAI, 5091 BALI,  
 GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys  
 5102 GGCAGGGCTGCATATGTGGCAAGTACCTCTTCAACTGGCAGTAAGAACAAAGCTAAA  
 CCGTCCCGACGGTATACACCGTTATGGAGAAGTTGACCCGTATTCTGTTGAGTT  
 ^  
 5113 NDEI,  
 LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr  
 5162 CTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTCACGGCTGGCTAC  
 GAGTGAGGTTATCGCCGGCGACCGGTCGACCTGAACAGGCCGACCAAGTGCCGACCGATG  
 ^ ^ ^  
 5174 NOTI, 5175 EAG1 XMA3, 5182 BALI, 5186 PVU2,  
 SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys  
 5222 AGCGGGGAGACATTATCACAGCGTGTCTCATGCCGGCCCGCTGGATCTGGTTTGC  
 TCGCCCCCTCTGTAAATAAGTGTGACAGAGTACGGGCCGGCGACCTAGACCAAAACG  
 ^  
 5240 DRA3,  
 LeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgMetSerThrAsn  
 5282 CTACTCCTGCTGCTGCAGGGTAGGCATCTACCTCCTCCCCAACGAATGAGCACGAAT  
 GATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGTTGGCTTACTCGTGCTTA  
 ^  
 5295 PSTI,  
 ProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGlnAspValLysPhe  
 5342 CCTAAACCTCAAAGAAAGACCAACGTAACACCAACCGGGCCGCAGGACGTCAAGTTC  
 GGATTGGAGTTCTTCTGGTTGCATTGTGGTTGGCCGCCGGCGTCTGCAGTTCAAG  
 ^ ^ ^  
 5380 NOTI, 5381 EAG1 XMA3, 5390 AAT2, 5401 SMA1 XMA1,  
 ProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeu  
 5402 CCGGGTGGCGGTCAAGATCGTTGGTAGTTACTTGCTGCAGGGCCCTAGATTG  
 GGCCCACCGCCAGTCTAGCAACCACCTCAAATGAACACGGCGTCCCCGGGATCTAAC  
 ^

## FIGURE 17 - Page 10

5449 APAI,

GlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGlyArgArgGlnPro  
5462 GGTGTGCGCGCGACGAGAAAGACTTCCGAGCGGTGCAACCTCGAGGTAGACGTAGCCT  
CCACACGCGCGCTGCTCTTCTGAAGGCTCGCCAGCGTTGGAGCTCCATCTGCAGTCGGA

5467 BSSH2, 5478 XMNI, 5502 XHOI, 5511 AAT2,

IleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpPro  
5522 ATCCCCAAGGCTCGTCGGCCCGAGGGCAGGACCTGGCTCAGCCCGGGTACCCCTGGCCC  
TAGGGGTTCCGAGCAGCCGGCTCCGCTGGACCCGAGTCGGGCCATGGGAACCAGGG

5548 ALWN1, 5558 ESP1, 5564 SMAI XMAI, 5568 KPNI,

LeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArg  
5582 CTCTATGGCAATGAGGGCTCGGGGTGGCGGGATGGCTCCTGTCTCCCCGTGGCTCTCGG  
GAGATACCGTTACTCCGACGCCACCCGCCCTACCGAGGACAGAGGGCACCGAGAGCC

ProSerTrpGlyProThrAspProArgArgSerArgAsnLeuGlyLysOC AM  
5642 CCTAGCTGGGGCCCCACAGACCCCCGGCGTAGGTGCGCAATTGGGTAAAGTAATAGTCG  
GGATCGACCCGGGGTGTCTGGGGCCGCATCCAGCGCGTAAACCCATTCAATTATCAGC

5650 APAI, 5698 SALI,

5702 AC  
TG

## FIGURE 18 - Page 1

MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn  
 2 AGCTTACAAAACAAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC  
 TCGAATGTTTGTACCGACGTACGTCGAGTCCCAGATATTCCACGATCATGAGTTG  
 ^ ^ ^  
 1 HIND3, 24 NDEI, 52 SCAI,  
  
 ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp  
 62 CCCTCTGTTGCTGCAACACTGGCTTGGTACATGTCCAAGGCTCATGGGATCGAT  
 GGGAGACAAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCAGCTA  
 ^  
 116 CLAI,  
  
 ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr  
 122 CCTAACATCAGGACCGGGGTGAGAACAAATTACCACTGGCAGCCCCATCACGTACTCCACC  
 GGATTGTAGTCCTGGCCCCACTCTGTTAATGGTGACCGTCGGGTAGTGCATGAGGTGG  
  
 TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys  
 182 TACGGCAAGTCCCTTGCCGACGGCGGGTGCTGGGGGGCGTTATGACATAATAATTGT  
 ATGCCGTTCAAGGAACGGCTGCCGCCACGAGCCCCCGCGAATACTGTATTATAAACAA  
  
 AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln  
 242 GACGAGTGCCACTCCACGGATGCCACATCCATCTGGCATTGGCAGTGTCCCTGACCAA  
 CTGCTCACGGTGAGGTGCCTACGGTAGGTAGAACCCGTAACCGTGACAGGAACGGTT  
  
 AlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal  
 302 GCAGAGACTGCGGGGGCGAGACTGGTTGTGCTGCCACCGCCACCCCTCCGGGCTCCGTC  
 CGTCTCTGACCCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCCGAGGCAG  
 ^  
 303 ALWN1,  
  
 ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe  
 362 ACTGTGCCCATCCCAACATCGAGGAGGTTGCTCTGTCCACCACCGGAGAGATCCCTTT  
 TGACACGGGGTAGGGTTGTAGCTCCTCAACGAGACAGGTGGTGGCCTCTAGGGAAAAA  
  
 TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis  
 422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGAGACATCTCATCTGTCT  
 ATGCCGTTCCGATAGGGGAGCTTCATTAGTCCCCCTCTGTAGAGTAGAACAGTA  
  
 SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal  
 482 TCAAAAGAAGTGCACGAACTCGCCGCAAAGCTGGTCGCATTGGCATTCAATGCCGTG  
 AGTTTCTTCTCACGCTGCTTGAGCGCGTTTCGACCAGCGTAACCGTAGTTACGGCAC  
  
 AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal  
 542 GCCTACTACCGCGGTCTTGACGTGTCCGTACATCCCGACCAGCGCGATGTTGTCGTG  
 CGGATGATGGCGCCAGAACCTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAGCAGCAC  
 ^ ^  
 550 SAC2, 560 DRD1,  
  
 AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn  
 602 GCAACCGATGCCCTCATGACCGGCTATAACCGGGCACTTCGACTCGGTGATAGACTGCAAT  
 CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA  
 ^  
 615 BSPH1,

## FIGURE 18 - Page 2

662 ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle  
 ACGTGTGTCACCCAGACAGTCGATTCAGCCTGACCCCTACCTCACCAATTGAGACAATC  
 TGCACACAGTGGGTCTGTCAGCTAAAGTCGGAAGTGGATGGTAAGTCTGTTAG

722 ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys  
 ACGCTCCCCCAAGATGCTGTCCTCCGCACCAACGTCGGGGCAGGACTGGCAGGGGGAAAG  
 TGCGAGGGGGTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCTGACCGTCCCCCTTC

782 ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer  
 CCAGGCATCTACAGATTGTCGGCACCAGGGAGCGCCCTCCGGCATGTTGACTCGTCC  
 GGTCCGTAGATGTCTAACACCGTGGCCCCCTCGCGGGAGGCCGTACAAGCTGAGCAGG

816 BGLI, 833 DRD1,

842 ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr  
 GTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTGGTATGAGCTCACGCCGCCGAGACT  
 CAGGAGACACTCACGATACTCGTCCGACACGAACCATACTCGAGTGCAGGGCGGCTCTGA

881 SACI,

902 ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu  
 ACAGTTAGGCTACGAGCGTACATGAACACCCGGGCTTCCGTGTGCCAGGACCATCTT  
 TGTCAATCCGATGCTCGCATGTACTTGTGGGGCCCGAAGGGCACACGGTCTGGTAGAA

931 SMAI XMAI,

962 GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln  
 GAATTTGGGAGGGCGTCTTACAGGCCTCACTCATATAGATGCCACTTCTATCCCAG  
 CTTAAAACCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAAGGTC

985 STUI,

1022 ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla  
 ACAAAAGCAGAGTGGGAGAACCTTCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT  
 TGTTCGTCTCACCCCTTTGGAAGGAATGGACCATCGCATGGTCGGCACACCGCA

1069 DRA3,

1082 ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys  
 AGGGCTCAAGCCCTCCCCCATCGTGGGACCAGATGTGGAAGTGGTATTGCTCAAG  
 TCCCAGAGTTCGGGGAGGGTAGCACCTGGTACACCTCACAAACTAACGGGAGTTC

1142 ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle  
 CCCACCCCTCCATGGGCCAACACCCCTGCTATACAGACTGGCGCTGTCAGAATGAAATC  
 GGGTGGGAGGTACCCGGTTGTGGGGACGATATGTCGACCCGCGACAAGTCTTACTTAG

1150 NCOI,

1202 ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal  
 ACCCTGACGCACCCAGTCACCAAAATACATCATGACATGCTGGCCGACCTGGAGGTC  
 TGGGACTGCGTGGGTCACTGGTTATGTAGTACTGTACGTACGCCGCTGGACCTCCAG

1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,

1262 ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaTyrCysLeu  
 GTCACGAGCACCTGGGTGCTCGTGGCGCGTCTGGCTGCTTGGCCCGTATTGCCTG

## FIGURE 18 - Page 3

CAGTGCTCGTGGACCCACGAGCAACCGCCGCAGGACCGACGAAACCGGCGCATAACGGAC

1322    SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle  
 TCAACAGGCTGCGTGGTCATAGTGGCAGGGTCGTCTTGTCCGGGAAGCCGGCAATCATA  
 AGTTGTCCGACGCACCAGTATCACCGTCCCAGCAGAACAGGCCCTCGGCCGTTAGTAT

1369 NAEI,

1382    ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu  
 CCTGACAGGGAAAGTCCTCTACCGAGAGTTGATGAGATGGAAGAGTGCTCTCAGCACTTA  
 GGACTGTCCCTCAGGAGATGGCTCTCAAGCTACTTACCTCTCACGAGAGTCGTGAAT

1385 DRD1,

1442    ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu  
 CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTCAAGCAGAACGGCCCTCGGCCTC  
 GGCATGTAGCTCGTCCCTACTACGAGCGGCTCGTCAAGTTCGTCTCCGGGAGCCGGAG

1502    LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln  
 CTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCTGCTGTCCAGACCAACTGGCAA  
 GACGTCTGGCGCAGGGCAGTCCGTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT

1502 PSTI, 1507 TTH3I,

1562    LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu  
 AAACTCGAGACCTTCTGGCGAAGCATATGTGAACTTCATCAGTGGGATACAATACTTG  
 TTTGAGCTCTGGAAGACCCGTTCGTATACACCTTGAAGTAGTCACCCTATGTTATGAAC

1565 XHOI, 1586 NDEI,

1622    AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla  
 CGGGGCTTGTCAACGCTGCCTGGTAACCCCGCATTGCTTCATTGATGGCTTTACAGCT  
 CGCCCGAACAGTTGCGACGGACCATTGGGCGGTAAACGAAGTAACTACCGAAAATGTCGA

1643 BSTE2, 1677 ALWN1 PVU2,

1682    AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp  
 GCTGTCACCAGCCCCTAAACCACTAGCCAAACCCCTCTTCAACATATTGGGGGGGTGG  
 CGACAGTGGTCGGGTGATTGGTGTACGGTTGGGAGGAGAAGTTGTATAACCCCCCACC

1742    ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla  
 GTGGCTGCCAGCTCGCCGCCCGGTGCGCTACTGCCTTGTTGGCGCTGGCTAGCT  
 CACCGACGGGTGAGCGGGGGCCACGGCGATGACGGAAACACCCCGGACCGAATCGA

1794 ESP1,

1802    GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr  
 GGCGCCGCCATCGGCAGTGGACTGGGAAAGGTCTCATAGACATCCTGAGGGTAT  
 CCGCGGGTAGCCGTACAACCTGACCCCTCCAGGAGTATCTGTAGGAACGTCCCATA

1802 KAS1 NARI,

1862    GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer  
 GGCGCGGGCGTGCGGGAGCTTGTGGCATTCAAGATCATGAGCGGTGAGGTCCCTCC  
 CCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTCTAGTACTCGCCACTCCAGGGAGG

1878 SAC1, 1899 BSPH1,

## FIGURE 18 - Page 4

1922 ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly  
 ACGGAGGACCTGGTCAATCTACTGCCGCCATCCTCTGCCCGGAGCCCTCGTAGTCGGC  
 TGCCTCCTGGACCAGTTAGATGACGGCGGTAGGAGAGCAGGAGCATCAGCCG  
 1928 TTH3I,  
 1982 ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp  
 GTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCGAGGGGGCAGTGCAGTGG  
 CACCAGACACGTCGTTATGACGCGGCCGTGCAACCGGGCCGCTCCCCGTCACGTCACC  
 2004 NAEI, 2017 SMAI XMAI,  
 2042 MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal  
 ATGAACCGGCTGATAGCCTCGCCTCCGGGGAACCATGTTCCCCACGCACACTACGTG  
 TACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGTGCGTGATGCAC  
 2067 SMAI XMAI, 2093 DRA3,  
 2102 ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln  
 CCGGAGAGCGATGCAGCTGCCCGCCTACTGCCATACTCAGCAGCCTCACTGTAACCCAG  
 GGCCTCTCGCTACGTCGACGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGTC  
 2115 PVU2, 2159 ALWN1,  
 2162 LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer  
 CTCCTGAGGCGACTGCACCAGTGGATAAGCTGGAGTGTACCAACTCCATGCTCCGGTTCC  
 GAGGACTCCGCTGACGTGGCACCTATTGAGCCTCACATGGTGAGGTACGAGGCCAAGG  
 2164 MST2, 2220 ECON1,  
 2222 TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu  
 TGGCTAACGGGACATCTGGACTGGATATGCGAGGTGTTGAGCGACTTTAACACCTGGCTA  
 ACCGATTCCCTGTAGACCTGACCTATACGCTCACAACTCGCTGAAATTCTGGACCGAT  
 2282 LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr  
 AAAGCTAACGCTCATGCCACAGCTGCCTGGATCCCCTTGTGTCCTGCCAGCGCGGGTAT  
 TTTCGATTGAGTACGGTGTGACGGACCCCTAGGGAAACACAGGACGGTCGCGCCCATA  
 2285 ESP1, 2300 PVU2, 2310 BAMHI,  
 2342 LysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAlaGluIle  
 AAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC  
 TTCCCCCAGACCGCTCCCTGCCGTAGTACGTGTGAGCGACGGTACACCTCGACTCTAG  
 2402 ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet  
 ACTGGACATGTCAAAACGGGACGATGAGGATCGTCGGCTTAGGACCTGCAGGAACATG  
 TGACCTGTACAGTTTGCCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCCTTGTAC  
 2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,  
 2462 TrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeuProAla  
 TGGAGTGGGACCTCCCCATTAATGCCTACACCACGGGCCCTGTACCCCCCTCTGCG  
 ACCTCACCCCTGGAAGGGGTAATTACGGATGTGGTGCCTGGGACATGGGGGAAGGACGC  
 2480 ASE1, 2497 APAI,

## FIGURE 18 - Page 5

2522 ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln  
 CCGAACTACACGTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATAAGGCAG  
 GGCTTGATGTGCAAGCGCGATACTCCCACAGACGTCTCCTATGCACCTCTATTCCGTC

^

2553 PST1,

2582 ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln  
 GTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCGTGCCAG  
 CACCCCTGAAGGTATGCACTGCCACTGTGACTGTTAGAATTACGGGCACGGTC

^

2594 DRA3,

2642 ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro  
 GTCCCCATCGCCCCAATTTCACAGAATTGGACGGGGTGCGCCACATAGGTTGCGCCC  
 CAGGGTAGCGGGCTAAAAAGTGTCTAACCTGCCACGGATGTATCCAAACGCGGG

^

2702 ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro  
 CCCTGCAAGCCCTTGCTCGGGAGGAGGTATCATTAGACTCCACGAATACCCG  
 GGGACGTTCGGGAACGACGCCCTCCATAGTAAGTCTCATCCTGAGGTGTTATGGGC

^

2757 HGIE2,

2762 ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu  
 TAGGGTTCGCAATTACCTTGCAGGCCAACCGACGTGGCGTGTGACGTCCATGCTC  
 CATCCAGCGTTAATGGAACGCTCGGCTTGGCCTGCACCGGACAACGTGAGGTACGAG

^

2809 AAT2,

2822 ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro  
 ACTGATCCCTCCCATATAACAGCAGAGGCCGGCGAACGGTGGCGAGGGATCACCC  
 TGACTAGGGAGGGTATATTGTCGTCTCCGCCGGCCGCTTCCAACCGCTCCCTAGTGGG

^

2850 EAG1 XMA3,

2882 ProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAlaThrCys  
 CCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCAAGGCAACTTGC  
 GGGAGACACCGGTCGAGGAGCCGATCGGTGATAGGCAGGTAGAGAGTTCCGTTGAACG

^

^

2889 BALI, 2903 NHEI,

2942 ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln  
 ACCGCTAACCATGACTCCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAG  
 TGGCGATTGGTACTGAGGGACTACGACTCGAGTATCTCCGGTGGAGGATAACCTCCGTC

^

^

2966 ESP1, 2969 SACI,

3002 GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer  
 GAGATGGGCGGCAACATCACCAAGGGTTGAGTCAGAAAAGTGGTATTCTGGACTCC  
 CTCTACCCGCCGTGTAGTGGTCCCAACTCAGTCTTTGTTTCAACCACGACCTGAGG

^

3062 PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu  
 TTCGATCCGCTTGTGGCGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCCTG  
 AAGCTAGGCGAACACCGCCTCCTCGCTGCCCTCTAGAGGCATGGCGTCTTAGGAC

^

3096 BGL2,

^

ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro

## FIGURE 18 - Page 6

3122 CGGAAGTCTGGAGATTGCCAGGCCCTGCCGTTGGCGCGCCGGACTATAACCC  
 GCCTCAGAGCCTAAGGGTCCGGACGGCAAACCGCGCCGGCTGATATTGGG  
 ^ ^  
 3143 ALWN1, 3164 EAG1 XMA3,  
 ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro  
 3182 CCGCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACCACCTGTGGTCCATGGCTGCCCG  
 GGCATCACCTCTGCACCTTCGGCTGATGCTTGGTGGACACCAGGTACCGACGGC  
 ^ ^  
 3217 HGIE2, 3229 NCOI,  
 LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu  
 3242 CTTCCACCTCCAAAGTCCCTCCTGTGCCTCCGCCCTGGAAAGAACGGGACGGTGGCTC  
 GAAGGTGGAGGTTTCAGGGAGGACACGGAGGCGGAGCCTCTCGCCTGCCACCAGGAG  
 ^ ^  
 ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer  
 3302 ACTGAATCAACCTATCTACTGCCTGGCCGAGCTGCCACCAGAACGCTTGGCAGCTCC  
 TGACTTAGTTGGATAGATGACGAAACGGCTCGAGCGGTGGCTTCGAAACCGTCGAGG  
 ^ ^  
 3332 SACI, 3346 HIND3,  
 SerThrSerGlyIleThrGlyAspAsnThrThrSerSerGluProAlaProSerGly  
 3362 TCAACTCCGGCATTACGGCGACAATACGACAACATCCTCTGAGCCGCCCTCTGGC  
 AGTTGAAGGCCGTAATGCCGCTGTTATGCTGTGAGGAGACTCGGGCGGGAAAGACCG  
 ^ ^  
 CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro  
 3422 TGCCCCCCCCGACTCCGACGCTGAGTCCTATTCCCTCATGCCCGCCCTGGAGGGGAGCCT  
 ACGGGGGGCTGAGGCTCGACTCAGGATAAGGAGGTACGGGGGGACCTCCCCCTCGGA  
 ^  
 3437 EAM11051,  
 GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu  
 3482 GGGGATCCGGATCTTAGCGACGGGTATGGTCAACGGTCAGTAGTGAGGCCAACCGGGAG  
 CCCCTAGGCCTAGAATCGCTGCCAGTACCAAGTGGCCAGTCATCACTCCGGTTGCGCCTC  
 ^ ^  
 3484 BAMHI, 3485 BSAB1, 3487 BSPE1,  
 AspValValCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla  
 3542 GATGTCGTGTGCTCAATGTCTACTCTGGACAGGCGACTCGTACCCGTGCGCC  
 CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGGCAGTGGGCACGCGG  
 ^ ^  
 3589 DRA3, 3600 SAC2,  
 AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn  
 3602 GCGGAAGAACAGAAACTGCCCATCAATGCACTAAGCAACTCGTGTGCTACGTACCAACAT  
 CGCCTTCTGTCTTGACGGTAGTTACGTGATTGAGCAACGATGCAGTGGTGTAA  
 ^  
 3611 ALWN1, 3655 PFLM1,  
 LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp  
 3662 TTGGTGTATTCCACCACCTCACGCACTGCTTGCACAGGAGAACAGTACATTGAC  
 AACACACATAAGGTGGAGTGCACGAAACGGTTCCGTCTTCAGTGTAAACTG  
 ^  
 3681 DRA3,  
 ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla  
 3722 AGACTGCAAGTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTAAAGCAGCGCG

## FIGURE 18 - Page 7

TCTGACGTTCAAGACCTGTCGGTAATGGTCCTGCATGAGTCCTCCAATTCTCGTCGCCGC  
 SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis  
 3782 TCAAAAGTGAAGGCTAACTGCTATCCGTAGAGGAAGCTTGCAGCCTGACGCCACAC  
 AGTTTCACTCCGATTGAACGATAGGCATCTCCTCGAACGTCGGACTGCGGGGGTGTG  
 3816 HIND3,  
 SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla  
 3842 TCAGCCAATCCAAGTTGGTTATGGGGCAAAAGACGTCGGCTGCCATGCCAGAAAGGCC  
 AGTCGGTTAGGTTCAAACCAATACCCCGTTCTGCAGGCAACGGTACGGTCTTCGG  
 3875 AAT2, 3890 BGLI,  
 ValThrHisIleAsnSerValTrpLysAspLeuLeuAspAsnValThrProIleAsp  
 3902 GTAAACCCACATCAACTCCGTGGAAAGACCTCTGGAAAGACAATGTAACACCAATAGAC  
 CATTGGGTGAGTTGAGGCACACCTTCTGGAAAGACCTCTGTTACATTGTGGTTATCTG  
 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyArgLys  
 3962 ACTACCACATGGCTAAAGAACGAGGTTTCTGCGTTAGCCTGAGAAGGGGGTCGTAAG  
 TGATGGTAGTACCGATTCTGCTCCAAAGACGCAAGTCGGACTCTCCCCCAGCATTG  
 ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu  
 4022 CCAGCTCGTCTCATCGTGGTCCCGATCTGGCGTGCCTGAGAAGGGGGTCGTTG  
 GGTGAGCAGAGTAGCACAAAGGGCTAGACCCGCACCGCACACGCTTCTACCGAAAC  
 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr  
 4082 TACGACGTGGTACAAAGCTCCCTGGCGTGAATGGGAAGCTCCTACGGATTCCAATAC  
 ATGCTGCACCAATGTTGAGGGGAACCGGCACTACCCCTCGAGGATGCCAAGGTTATG  
 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysThrProMet  
 4142 TCACCAAGGACAGCGGGTTGAATTCTCGTCAAGCGTGGAACTCCAAGAAAACCCAAATG  
 AGTGGTCCTGTCGCCAACCTAACGGACACGTTGCACCTCAGGTTTTGGGGTTAC  
 4160 ECORI,  
 GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr  
 4202 GGGTTCTCGTATGATACCCGCTGTTGACTCCACAGTCACTGAGAGCGACATCCGTACG  
 CCCAAGAGCATACTATGGGCGACGAAACTGAGGTGTCAGTGAATCTCGCTGTAGGCATGC  
 4229 DRD1, 4236 ALWN1,  
 GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer  
 4262 GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCCAAGCCCGTGGCCATCAAGTCC  
 CTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGTTGGGCGCACCGGTAGTCAGG  
 4301 BGLI, 4308 BALI,  
 LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly  
 4322 CTCACCGAGAGGCTTATGTTGGGGCCCTCTTACCAATTCAAGGGGGAGAACTGCGGC  
 GAGTGGCTCTCGAAATACAACCCCGGGAGAATGGTTAACACTGGAGCTGGGGTTGGGCGCACCGGTAGTCAGG  
 4345 APAI,  
 TyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCys  
 4382 TATCGCAGGTGCCGCGCAGCGCGTACTGACAACTAGCTGTGGTAACACCCCTCACTTGC  
 ATAGCGTCCACGGCGCCTCGCCGATGACTGTTGATCGACACCATTGTGGGAGTGAACG

## FIGURE 18 - Page 8

4442 TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal  
 TACATCAAGGCCCGGGCAGCCTGTCGAGCCGCAGGGCTCCAGGACTGCACCATGCTCGTG  
 ATGTAGTTCCGGGCCCCGTCGGACAGCTCGGCGTCCCGAGGTCTGACGTGGTACGAGCAC  
 ^  
 4452 SMAI XMAI,

4502 CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer  
 TGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGTCCAGGAGGACGCGCGAGC  
 ACACCGCTGCTGAATCAGCAATAGACACTTTCGCCTCCAGGTCTCGGCCGCTCG  
 ^ ^  
 4508 DRD1, 4511 TTH3I,

4562 LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln  
 CTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCTGGGACCCCCCACAA  
 GACTCTCGGAAGTGCCTCCGATACTGGTCCATGAGGCGGGGGGACCCCTGGGGGTGTT

4622 ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp  
 CCAGAATACGACTTGGAGCTATAACATCATGCTCCTCAAACGTGTCAGTCGCCACGAC  
 GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGGGTGCT  
 ^  
 4637 SACI,

4682 GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla  
 GGCGCTGGAAAGAGGGTCTACTACCTCACCGTGACCCCTACAACCCCCCTCGCGAGAGCT  
 CCGCGACCTTCTCCCAGATGATGGAGTGGGACTGGGATGTTGGGGAGCGCTCTCGA  
 ^  
 4731 NRUI,

4742 AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe  
 GCGTGGGAGACAGCAAGACACACTCCAGTCAATTCTGGCTAGGCAACATAATCATGTT  
 CGCACCCCTCTGCGTTCTGTGAGGTCAAGGACCGATCCGTTGTATTAGTACAAA

4802 AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla  
 GCCCCCACACTGTGGCGAGGATGATACTGATGACCCATTCTTAGCGTCCTTATAGCC  
 CGGGGTGTGACACCCGCTCTACTATGACTACTGGTAAAGAAATCGCAGGAATATCGG  
 ^ ^  
 4806 PFLM1, 4807 DRA3,

4862 ArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu  
 AGGGACCAGCTGAACAGGCCCTCGATTGCGAGATCTACGGGCCTGCTACTCCATAGAA  
 TCCCTGGTCGAACTTGTCCGGGAGCTAACGCTCTAGATGCCCGGACGATGAGGTATCTT  
 ^  
 4893 BGL2,

4922 ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis  
 CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTCACCCAC  
 GGTGACCTAGATGGAGGTTAGTAAGTTCTGAGGTACCGGAGTCGCGTAAAGTGAGGTG  
 ^  
 4954 NCOI,

4982 SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro  
 AGTTACTCTCCAGGTGAAATCAATAGGTGGCCGCATGCCTCAGAAAACCTGGGGTACCG  
 TCAATGAGAGGTCCACTTAGTTATCCCACCGGGTACGGAGTCTTTGAACCCCATGGC  
 ^ ^  
 5015 SPHI, 5035 KPNI,

## FIGURE 18 - Page 9

5042 ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly  
CCCTTCGAGCTTGGAGACACCAGGGCCGGAGCGTCCCGCCTAGGCTCTGGCCAGAGGA  
GGGAACGCTCGAACCTCTGTGGCCGGCCTCGCAGGCGCGATCCGAAGACCGGTCTCCT  
5064 APAI, 5091 BALI,  
5102 GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys  
GGCAGGGCTGCCATATGTGGCAAGTACCTCTCAAACGGCAGTAAGAACAAAGCTCAA  
CCGTCCCCACGGTATAACACCGTTATGGAGAAGTTGACCCGTCAATTCTTGGAGTT  
5113 NDEI,  
5162 LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr  
CTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTACGGCTGGCTAC  
GAGTGAGGTTATCGCCGGCGACCGGTCGACCTGAACAGGCCGACCAAGTGCCGACCGATG  
5174 NOTI, 5175 EAG1 XMA3, 5182 BALI, 5186 PVU2,  
5222 SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys  
AGCGGGGAGACATTTATCACAGCGTCTCATGCCCGCCCGCTGGATCTGGTTTGC  
TCGCCCTCTGTAAATAGTGTGCACAGAGTACGGCCGGCGACCTAGACCAAAACG  
5240 DRA3,  
5282 LeuLeuLeuLeuAlaAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgMetSerThrAsn  
CTACTCCTGCTGCTGCAGGGTAGGCATCTACCTCCTCCCAACCGAATGAGCACGAAT  
GATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGTTGGCTTACTCGTCTTA  
5295 PSTI,  
5342 ProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGlnAspValLysPhe  
CCTAACCTCAAAGAAAGACCAACGTAACACCAACCGGGCCGCAGGACGTCAAGTTC  
GGATTTGGAGTTCTTCTGGTTGCATTGTGGTTGGCCGGCGTCCGTCAAG  
5380 NOTI, 5381 EAG1 XMA3, 5390 AAT2, 5401 SMA1 XMA1,  
5402 ProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeu  
CCGGGTGGCGGTCAAGATCGTGGAGTTACTTGTGGCCGCAGGGCCCTAGATTG  
GCCACCCACCGCCAGTCTAGCAACCACCTCAAATGAACAAACGGCGCGTCCCCGGGATCTAAC  
5449 APAI,  
5462 GlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGlyArgArgGlnPro  
GGTGTGCGCGCAGAGAAAGACTTCCGAGCGGTGCGAACCTCGAGGTAGACGTAGCCT  
CCACACGCGCGCTGCTCTGAAGGCTGCCAGCGTTGGAGCTCCATCTGCAGTCGGA  
5467 BSSH2, 5478 XMNI, 5502 XHOI, 5511 AAT2,  
5522 IleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpPro  
ATCCCCAAGGCTCGTCGGCCCGAGGGCAGGACCTGGGCTCAGCCGGTACCCCTGGCC  
TAGGGGTTCCGAGCAGCCGGCTCCCGTCTGGACCCGAGTCGGGCCATGGGAACCGGG  
5548 ALWN1, 5558 ESP1, 5564 SMA1 XMA1, 5568 KPN1,  
5582 LeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArg  
CTCTATGGCAATGAGGGCTCGGGGTGGCGGGATGGCTCCTGTCTCCCGTGGCTCTCGG  
GAGATACCGTTACTCCGACGCCACCCGCCCTACCGAGGACAGAGGGCACCGAGAGCC

## FIGURE 18 - Page 10

5642 ProSerTrpGlyProThrAspProArgArgArgSerArgAsnLeuGlyLysValIleAsp  
CCTAGCTGGGGCCCCACAGACCCCCGGCGTAGGTCGCGCAATTGGGTAAAGGTATCGAT  
GGATCGACCCCGGGGTGTCTGGGGCCGCATCCAGCGCGTTAAACCCATTCCAGTAGCTA  
5650 APAI, 5696 CLAI,  
  
5702 ThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProLeuValGlyAlaProLeu  
ACCCTTACGTGCGGCTTCGCCGACCTCATGGGTACATACCGCTCGTCGGCGCCCTCTT  
TGGGAATGCACGCCGAAGCGGCTGGAGTACCCATGTATGGCGAGCAGCCGCGGGGAGAA  
5724 HGIE2, 5750 KAS1 NARI, 5756 ECON1,  
  
5762 GlyGlyAlaAlaArgAlaLeuAlaHisGlyValArgValLeuGluAspGlyValAsnTyr  
GGAGGCCTGCCAGGGCCCTGGCGCATGGCGTCCGGTTCTGGAAGACGGCGTGAACATAT  
CCTCCCGCACGGTCCCAGGACCGCGTACCGCAGGCCAAGACCTTCTGCCGCACTTGATA  
5772 BSTXI, 5775 APAI,  
  
5822 AlaThrGlyAsnLeuProGlyCysSerOC AM  
GCAACAGGGAACCTCCTGGTTGCTCTTAATAGTCGAC  
CGTTGTCCTGGAAAGGACCAACGAGAATTATCAGCTG  
5854 SALI,

FIGURE 19

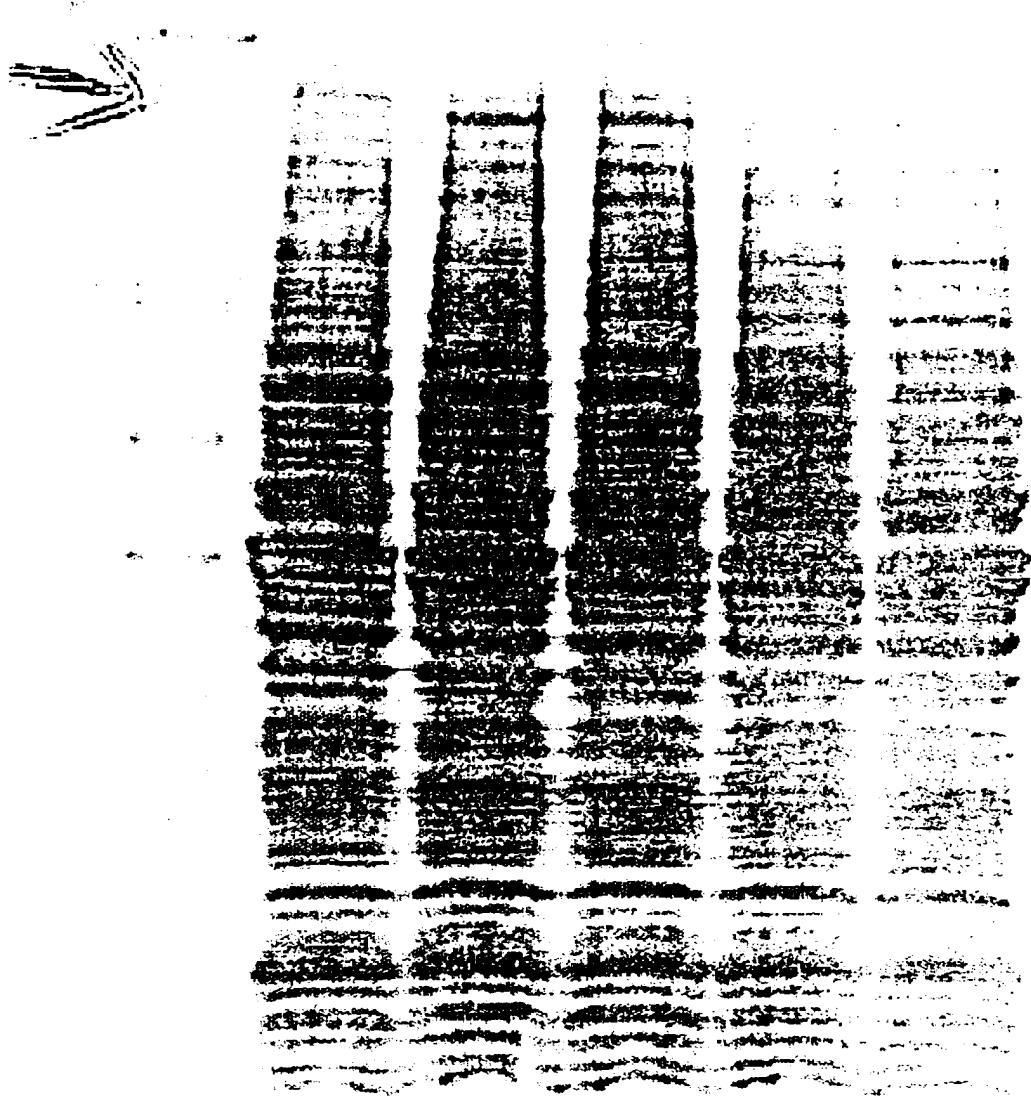


FIGURE 20 - Page

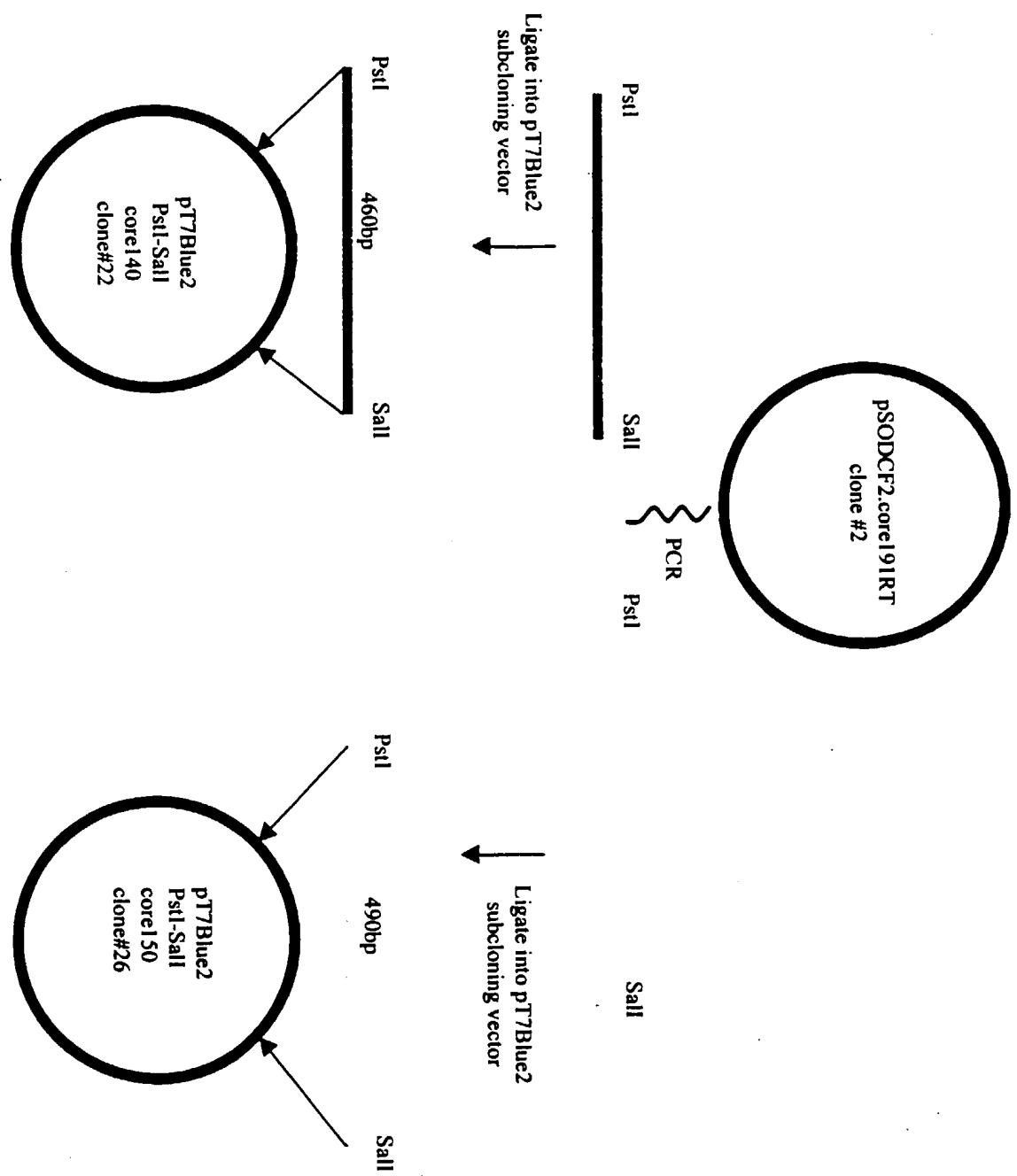
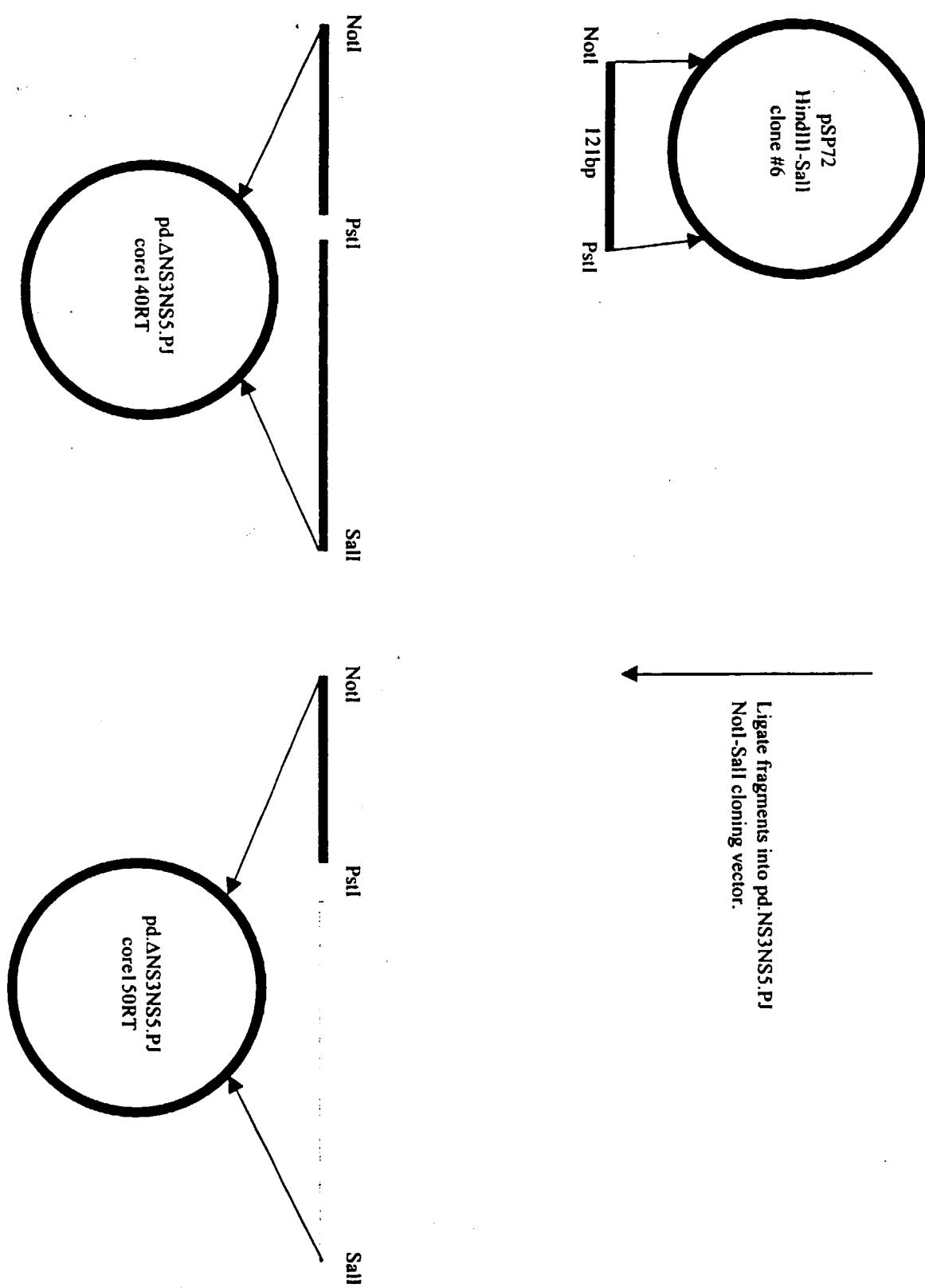


FIGURE 20 - Pa



## FIGURE 21 - Page 1

Met Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn  
2 AGCTTACAAAACAAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC  
TCGAATGTTTGTGTTACCGACGTACGTCGAGTCCCGATATTCCACGATCATGAGTTG  
^ ^ ^  
1 HIND3, 24 NDEI, 52 SCAI,  
Pro Ser Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr Met Ser Lys Ala His Gly Ile Asp  
62 CCTCTGTTGCTGCAACACTGGGCTTGGTGCTTACATGTCACAGGCTCATGGGATCGAT  
GGGAGACAAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCCTAGCTA  
^  
116 CLAI,  
Pro Asn Ile Arg Thr Gly Val Arg Thr Ile Thr Thr Gly Ser Pro Ile Thr Tyr Ser Thr  
122 CCTAACATCAGGACCGGGGTGAGAACAAATTACCACTGGCAGCCCCATACGTACTCCACC  
GGATTGTAGTCTGGCCCCACTCTTGTAAATGGTGACCGTCGGGTAGTGCATGAGGTGG  
Tyr Gly Lys Phe Leu Ala Asp Gly Gly Cys Ser Gly Gly Ala Tyr Asp Ile Ile Ile Cys  
182 TACGGCAAGTTCTTGCCACGGCGGGTGCTCGGGGGCGCTTATGACATAATAATTGT  
ATGCCGTTCAAGGAACGGCTGCCGCCACGAGCCCCCGCGAATACTGTATTATAAACAA  
Asp Glu Cys His Ser Thr Asp Ala Thr Ser Ile Leu Gly Ile Gly Thr Val Leu Asp Gln  
242 GACGAGTGCCTACTCCACGGATGCCACATCCATCTTGGCATTGGCAGTGTCCCTTGACCAA  
CTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCGTAACCGTACAGGAACCTGGTT  
Ala Glu Thr Ala Gly Ala Arg Leu Val Val Leu Ala Thr Ala Thr Pro Pro Gly Ser Val  
302 GCAGAGACTGCGGGGGCGAGACTGGTTGTGCTGCCACGCCACCCCTCCGGCTCCGTC  
CGTCTCTGACGCCCGCTCTGACCAACACGAGCGGTGGCGTGGGGAGGCCCGAGGCAG  
^  
303 ALWN1,  
Thr Val Pro His Pro Asn Ile Glu Glu Val Ala Leu Ser Thr Thr Gly Glu Ile Pro Phe  
362 ACTGTGCCCATCCAACATCGAGGAGGTTGCTCTGTCACCACCGGAGAGATCCCTTT  
TGACACGGGGTAGGGTTGTAGCTCCACGAGACAGGTGGTGGCCTCTAGGGAAAA  
Tyr Gly Lys Ala Ile Pro Leu Glu Val Ile Lys Gly Gly Arg His Leu Ile Phe Cys His  
422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGAGACATCTCATCTCTGTCT  
ATGCCGTTCCGATAGGGGGAGCTCATTAGTTCCCCCTCTGTAGAGTAGAACAGTA

**FIGURE 21 - Page 2**

482 SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal  
 TCAAAGAAGAAGTGCACGAACTGCCGAAAGCTGGTCGATTGGGCATCAATGCCGTG  
 AGTTTCTCTTCACGCTGCTTGAGCGCGTTCGACCAGCGTAACCCGTAGTTACGGCAC  
 AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal  
 542 GCCTACTACCGCGGTCTTGACGTGTCGTACATCCCGACCAGCGCGATGTTGTCGTGCGTGC  
 CGGATGATGGCGCCAGAACCTGCACAGGCAGTAGGGCTGGTCGCCGTACAACAGCAC  
 550 SAC2, 560 DRD1,  
 AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn  
 602 GCAACCAGATGCCCTCATGACCGGCTATACCGGCACCTCGACTCGGTGATAGACTGCAAT  
 CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA  
 615 BSPH1,  
 ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle  
 662 ACGTGTCACCCAGACAGTCGATTCAGCCTTGACCCCTACCTTCACCATGAGACAATC  
 TGCACACAGTGGTCTGTCAGCTAAAGTCGAACTGGGATGGAAGTGGTAACCTGTTAG  
 ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys  
 722 ACGCTCCCCAAGATGCTCTCCGCCTCAACGTCGGGCAGGACTGGCAGGGGGAAAG  
 TGCGAGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCCTGACCGTCCCCCTTC  
 ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer  
 782 CCAGGCATCTACAGATTGTGGCACCGGGGAGCGCCCCCTCCGGCATGTTGACTCGTCC  
 GGTCCGTAGATGTCTAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGG  
 816 BGLI, 833 DRD1,  
 ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr  
 842 GTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTGGTATGAGCTCACGCCCGCAGACT  
 CAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAATGCGGGCGCTCTGA  
 881 SAC1,  
 ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu  
 902 ACAGTTAGGCTACGAGCGTACATGAACACCCGGGCTTCCGTGTGCCAGGACCATCTT  
 TGTCATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCCTGGTAGAA  
 931 SMA1 XMA1,  
 GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln  
 962 GAATTTGGGAGGGCGTCTTACAGGCCTCACTCATATAGATGCCACTTCTATCCAG  
 CTTAAACCCCTCCCGAGAAATGTCCGGAGTGAGTATCTACGGGTGAAAGATAGGGTC  
 985 STUI,  
 ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla  
 1022 ACAAAAGCAGAGTGGGGAGAACCTCCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT  
 TGTTCGTCTCACCCCTTGGAGGAATGGACCATCGCATGGTTGGCACACGCGA  
 1069 DRA3,  
 ArgAlaGlnAlaProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys  
 1082 AGGGCTCAAGCCCCCTCCCCCATCGTGGACCAGATGTGGAAGTGTGTTGATTGCCCTCAAG

# FIGURE 21 - Page 3

TCCCGAGTTGGGGAGGGGGTAGCACCCCTGGTCTACACCTTCACAAACTAAGCGGAGTC

1142 ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle  
 CCCACCCCTCCATGGGCCAACACCCCTGCTATACAGACTGGCGCTGTTCAGAATGAAATC  
 GGGTGGGAGGTACCCGGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAG  
 ^  
 1150 NCOI,

1202 ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal  
 ACCCTGACGCACCCAGTCACCAAATACATCATGACATGCGATGTCGGCCGACCTGGAGGTC  
 TGGGACTGCGTGGGTCACTGGTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG  
 ^ ^ ^ ^  
 1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,

1262 ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu  
 GTCACAGGACACTGGGTGCTCGTTGGCGCGTCTGGCTGCTTGGCCGCGTATTGCCTG  
 CAGTGCTCGTGGACCCACGAGCAACCGCCGCAGGACCGACGAAACCGGCATAACGGAC

1322 SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle  
 TCAACAGGCTGCGTGGTCATAGTGGCAGGGTCGTCTTGTCCGGGAAGCCGGCAATCATA  
 AGTTGTCCGACGCACCAGTATCACCCGTCCCAGCAGAACAGGCCCTCGGCCGTTAGTAT  
 ^  
 1369 NAEI,

1382 ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu  
 CCTGACAGGGAAAGTCCTCTACCGAGAGATCGATGAGATGGAAGAGTGCCTCTCAGCACTTA  
 GGACTGTCCCTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCACGAGAGTCGTGAAT  
 ^  
 1385 DRD1,

1442 ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu  
 CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTGGCCTC  
 GGCATGTAGCTCGTCCCTACTACGAGCGGCTCGTCAAGTTCGTCTCCGGGAGCCGGAG

1502 LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln  
 CTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCTGCTGTCCAGACCAACTGGCAA  
 GACGTCTGGCGCAGGGCAGTCCGTCTCAAATAGCGGGGACGACAGGTCTGGTTGACCGTT  
 ^ ^  
 1502 PSTI, 1507 TTH3I,

1562 LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu  
 AAACCTCGAGACCTCTGGCGAAGCATATGTGGAACATTCATCAGTGGGATAACAATACTTG  
 TTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCCTATGTTATGAAC  
 ^  
 1565 XHOI, 1586 NDEI,

1622 AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla  
 GCGGGCTTGTCAACGCTGCCTGGTAACCCGCCATTGCTTCATTGATGGCTTACAGCT  
 CGCCCGAACAGTTGCGACGGACCATTGGGGCGGTAAACGAAGTAACCGAAAATGTCGA  
 ^  
 1643 BSTE2, 1677 ALWN1 PVU2,

1682 AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp  
 GCTGTCACCAGCCCCACTAACCACTAGCCAAACCCCTCCTCTCAACATATTGGGGGGTGG  
 CGACAGTGGTCGGGTGATTGGTATCGGTTGGGAGGAGTTGTATAACCCCCCACC

**FIGURE 21 - Page 4**

1742 ValAlaAlaGlnLeuAlaAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla  
 GTGGCTGCCAGCTGCCGCCCCGGTGCCTACTGCCCTTGTGGCGCTGGCTTAGCT  
 CACCGACGGGTCGAGCGGGGGGCCACGGCGATGACGGAAACACCCCGGACCGAATCGA

1794 ESP1,

1802 GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr  
 GCGCCGCCATCGGCAGTGGACTGGGAAGGTCTCATAGACATCCTTGCAGGGTAT  
 CCGCGGGTAGCGTCACAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCCATA

1802 KAS1 NARI,

1862 GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer  
 GGCGCGGGCGTGGCGGGAGCTTGTGGATTCAAGATCATGAGCGGTGAGGTCCCCTCC  
 CCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTGCCACTCCAGGGAGG

1878 SAC1, 1899 BSPH1,

1922 ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly  
 ACGGAGGACCTGGTCAATCTACTGCCGCCATCCTCTGCCCGGAGCCCTCGTAGTCGGC  
 TGCCCTCTGGACCACTAGATGACGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG

1928 TTH3I,

1982 ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp  
 GTGGTCTGTGCAGCAATACTGCAGCTGGCCATCCTCTGCCCGGAGGGGGCAGTGCAGTGG  
 CACCAAGACACGTCGTTATGACGCGGCCGTGCAACCGGGCCGCTCCCCCGTCACGTCACC

2004 NAEI, 2017 SMAI XMAI,

2042 MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal  
 ATGAACCGGCTGATAGCCTTCGCCTCCGGGGAAACATGTTCCCCCACGCACACTACGTG  
 TACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTGGTACAAAGGGGGTGCCTGATGCAC

2067 SMAI XMAI, 2093 DRA3,

2102 ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln  
 CCGGAGAGCGATGCAGCTGCCGCGTCACTGCCATACTCAGCAGCCTCACTGTAACCCAG  
 GGCCTCTCGCTACGTCGACGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGGTC

2115 PVU2, 2159 ALWN1,

2162 LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer  
 CTCCTGAGGCAGTCACAGTGGATAAGCTGGAGTGTACCACTCCATGCTCCGGTTCC  
 GAGGACTCCGCTGACGTGGTCACCTATTGAGCCTCACATGGTGGAGGTACGAGGCCAAGG

2164 MST2, 2220 ECON1,

2222 TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu  
 TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAACGACTGGCTA  
 ACCGATTCCCTGTAGACCCCTGACCTATACGCTCCACAACTCGCTGAAATTCTGGACCGAT

2282 LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr  
 AAAGCTAACGCTCATGCCACAGCTGCCCTGGATCCCTTGTGTCCCTGCCAGCGGGTAT  
 TTTCGATTGAGTACGGTGTGACGGACCCCTAGGGAAACACAGGACGGTCGCGCCCCATA

2285 ESP1, 2300 PVU2, 2310 BAMHI,

## FIGURE 21 - Page 5

2342 LysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAlaGluIle  
 AAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC  
 TTCCCCCAGACCGCTCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAG

2402 ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet  
 ACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCTAGGACCTGCAGGAACATG  
 TGACCTGTACAGTTTGCCTGCTACTCCTAGCAGCCAGGATCTGGACGTCTTGAC  
 ^ ^ ^  
 2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,

2462 TrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeuProAla  
 TGGAGTGGGACCTCCCCATTAAATGCCACACCCACGGGCCCCGTACCCCCCTTCCTGCG  
 ACCTCACCCCTGGAAGGGTAATTACGGATGTGGTCCCCGGGACATGGGGGAAGGACGC  
 ^  
 2480 ASE1, 2497 APAI,

2522 ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln  
 CCGAACTACACGTTCGCGTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATAAGGCAG  
 GGCTTGATGTGCAAGCGCGATACCTCCCACAGACGTCTCCTATGCACCTCTATTCCGTC  
 ^  
 2553 PSTI,

2582 ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln  
 GTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTAAATGCCGTGCCAG  
 CACCCCCCTGAAGGTGATGCACTGCCCATACTGATGACTGTTAGAATTACGGCACGGTC  
 ^  
 2594 DRA3,

2642 ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro  
 GTCCCCATGCCCGAATTTTACAGAAATTGGACGGGGTGCCTACATAGGTTGCGCCC  
 CAGGGTAGCGGGCTTAAAAAGTGTCTAACCTGCCAACGCGGATGTATCCAAACGCGGG

2702 ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro  
 CCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTAGAGTAGGACTCCACGAATACCCG  
 GGGACGTTCGGGAACGACGCCCTCCTCCATAGTAAGTCTCATCCTGAGGTGCTTATGGC  
 ^  
 2757 HGIE2,

2762 ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu  
 GTAGGGTCGCAATTACCTTGCAGGCCGAACCGGACGTGGCGTGTGACGTCCATGCTC  
 CATCCCAGCGTTAATGGAACGCTCGGGCTTGGCCTGCACCGGACAACACTGCAGGTACGAG  
 ^  
 2809 AAT2,

2822 ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro  
 ACTGATCCCTCCCATATAACAGCAGAGGCAGCCGGCGAAGGTTGGCGAGGGGATCACC  
 TGACTAGGGAGGGTATATTGTCGTCTCCGCCGGCCGCTTCCAACCGCTCCCTAGTGGG  
 ^  
 2850 EAG1 XMA3,

2882 ProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAlaThrCys  
 CCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCAACTTGC  
 GGGAGACACCGGTCGAGGAGCCGATCGGTCGATAGGCGAGGTAGAGAGTTCCGTTGAACG  
 ^ ^  
 2889 BALI, 2903 NHEI,

## FIGURE 21 - Page 6

2942 ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln  
 ACCGCTAACCATGACTCCCCCTGATGCTGAGCTCATAGAGGCCAACCTCTATGGAGGCAG  
 TGGCGATTGGTACTGAGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACCTCCGTC  
 ^ ^  
 2966 ESP1, 2969 SACI,  
 3002 GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer  
 GAGATGGGCGGCAACATCACCAAGGGTTGAGTCAGAAAACAAAGTGGTATTCTGGACTCC  
 CTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTGTTCAACCCTAACGACCTGAGG  
 3062 PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu  
 TTCGATCCGCTTGTGGCGGGAGGAGCAGCAGGAGATCTCCGTACCCGCAGAAATCCTG  
 AAGCTAGGCGAACACCGCCCTCCTCGCTGCCCTCTAGAGGCATGGCGTCTTAGGAC  
 ^  
 3096 BGL2,  
 3122 ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro  
 CGGAAGTCTCGGAGATTGCCAGGCCCTGCCGTTGGCGCGCCGGACTATAACCCC  
 GCCTTCAGAGCCTCTAACGGGGTCCGGGACGGGAAACCCGCGCCGGCTGATATTGGGG  
 ^ ^  
 3143 ALWN1, 3164 EAG1 XMA3,  
 3182 ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro  
 CCGCTAGTGGAGACGTGGAAAAAGCCGACTACGAACCACCTGTGGTCCATGGCTGCCG  
 GGCATCACCTCTGCACCTTCGGCTGATGCTGGTGGACACCAGGTACCGACGGG  
 ^ ^  
 3217 HGIE2, 3229 NCOI,  
 3242 LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu  
 CTTCCACCTCCTAAAGTCCCCCTCCTGTGCCTCCGCCCTCGGAAGAACGGGACGGTGGCCTC  
 GAAGGTGGAGGTTCAAGGGAGGACACGGAGGCGAGCCTTTCGCCTGCCACCAGGAG  
 3302 ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer  
 ACTGAATCAACCCTATCTACTGCCTGGCCGAGCTGCCACCAGAACGTTGGCAGCTCC  
 TGACTTAGTTGGATAGATGACGGAACCGGCTGAGCGGTGGCTTCGAAACCGTCGAGG  
 ^ ^  
 3332 SACI, 3346 HIND3,  
 3362 SerThrSerGlyIleThrGlyAspAsnThrThrSerSerGluProAlaProSerGly  
 TCAACTCCGGCATTACGGCGACAATACGACAAACATCCTCTGAGCCCCCCCCTCTGGC  
 AGTTGAAGGCCGTAATGCCGCTGTTATGCTGTTGAGGAGACTGGGGGGAAAGACCG  
 3422 CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro  
 TGCCCCCCCCGACTCCGACGCTGAGTCTATTCCCATGCCCTGGAGGGGAGCCT  
 ACGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGACCTCCCCCTCGGA  
 ^  
 3437 EAM11051,  
 3482 GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu  
 GGGGATCCGGATCTTAGCGACGGGTCATGGTCAACGGTCAGTAGTGAGGCCAACCGGGAG  
 CCCCTAGGCCTAGAATCGCTGCCAGTACCAAGTGGCCAGTCATCACTCCGGTTGCGCCTC  
 ^ ^ ^  
 3484 BAMHI, 3485 BSAB1, 3487 BSPE1,  
 3542 AspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla  
 GATGTCGTGTGCTGCTCAATGTCTTACTCTGGACAGGCGCACTCGTCACCCCGTGC  
 CTACAGCACACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGCACCGGG

# FIGURE 21 - Page

3589 DRA3, 3600 SAC2,

3602 AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn  
GCGGAAGAACAGAAACTGCCATCAATGCACTAAGCAACTCGTTGCTACGTACCCACAAT  
CGCCTTCTTGTCTTGACGGGTAGTTACGTGATTGAGCAACGATGCAGTGGTGTAA

3611 ALWN1, 3655 PFLM1,

3662 LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp  
TTGGTGTATTCCACCACCTCACGCAGTGCTGCCAAAGGCAGAAGAAAGTCACATTGAC  
AACCACATAAGGTGGTGGAGTGCACGAACGGTTCCGTCTTCAGTGTAAACTG

3681 DRA3,

3722 ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla  
AGACTGCAAGTTCTGGACAGCCATTACCAAGGACGTACTCAAGGAGGTTAAAGCAGCGCG  
TCTGACGTTCAAGACCTGTCGGAATGGTCTGCATGAGTTCCCTCCAATTCTCGCCGC

3782 SerLysValLysAlaAsnLeuLeuSerValGluAlaCysSerLeuThrProProHis  
TCAAAAGTGAAGGCTAACCTGCTATCCGTAGAGGAAGCTGACGCCCTGACGCCACAC  
AGTTTCACTTCCGATTGAACGATAGGCATCTCCTCGAACGTCGGACTGCGGGGGTGTG

3816 HIND3,

3842 SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla  
TCAGCCAAATCCAAGTTGGTTATGGGCAAAGACGTCGTTGCCATGCCAGAAAGGCC  
AGTCGGTTAGGTTCAAACCAATACCCGTTTCTGCAGGCAACGGTACGGTCTTCCGG

3875 AAT2, 3890 BGLI,

3902 ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp  
GTAACCCACATCAACTCCGTGTGGAAAGACCTCTGGAAAGACAATGTAACACCAATAGAC  
CATTGGGTGTAGTTGAGGCACACCTTCTGGAAAGACCTCTGTTACATTGTGGTTATCTG

3962 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyArgLys  
ACTACCACATGGCTAACGAGGTTCTGCCTCAGCCTGAGAAGGGGGTGTAAAG  
TGATGGTAGTACCGATTCTGCTCAAAGACGCAAGTCGGACTCTCCCCCAGCATTG

4022 ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu  
CCAGCTCGTCTCATCGTGTCCCCGATCTGGCGTGCGCGTGTGCGAAAAGATGGCTTG  
GGTCGAGCAGAGTAGCACAAGGGGCTAGACCCGCACCGCACACGCTTTCTACCGAAAC

4082 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr  
TACGACGTGGTTACAAAGCTCCCTGGCCGTATGGGAAGCTCTACGGATTCCAATAC  
ATGCTGCACCAATGTTCGAGGGGACCGGACTACCCCTCGAGGATGCCTAAGGTTATG

4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet  
TCACCAGGACAGCGGGTTGAATTCCCTCGTGCAGCGTGGAAAGTCCAAGAAAACCCAATG  
AGTGGTCCTGTCGCCAACTTAAGGAGCACGTTCGCACCTTCAGGTTCTTGGGTTAC

4160 ECORI,

4202 GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr  
GGGTTCTCGTATGATAACCGCTGCTTGACTCCACAGTCAGTGAGAGCGACATCCGTACG  
CCCAAGAGCATACTATGGCGACGAAACTGAGGTGTAGTGAATCTCGCTGTAGGCATGC

## FIGURE 21 - Page 8

4229 DRD1, 4236 ALWN1,

4262 GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer  
GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCCAAGCCCGTGGCCATCAAGTCC  
CTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGTTGGGGCGACCGGTAGTCAGG

4301 BGLI, 4308 BALI,

4322 LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly  
CTCACCGAGAGGGCTTATGTTGGGGCCCTCTTACCAATTCAAGGGGGAGAACTGCGGC  
GAGTGGCTCTCCGAAATACAACCCCCGGGAGAATGGTTAAGTTCCCCCTCTTGACGCCG

4345 APAI,

4382 TyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCys  
TATCGCAGGTGCCGCGCAGCGCGTACTGACAAGTAGCTGTGGTAACACCCTCACTTGC  
ATAGCGTCCACGGCGCGCTGCCGCATGACTGTGATCGACACCATTGTGGAGTGAACG

4442 TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal  
TACATCAAGGCCGGCAGCCTGTCGAGCCGCAGGGCTCCAGGACTGCACCATGCTCGTG  
ATGTAGTTCCGGGCCCCGTCGGACAGCTCGGCGTCCCAGGTCTGACGTGGTACGAGCAC

4452 SMAI XMAI,

4502 CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer  
TGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCAGGGCTCCAGGAGGACGCGCGAGC  
ACACCGCTGCTGAATCAGCAATAGACACTTCGCGCCCCCAGGTCTCCTGCGCCGCTCG

4508 DRD1, 4511 TTH3I,

4562 LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln  
CTGAGAGCCTCACGGAGGCTATGACCAGGTACTCCGCCCCCCTGGGACCCCCACAA  
GACTCTCGGAAGTGCCTCCGATACTGGTCCATGAGGCGGGGGGACCCCTGGGGGTGTT

4622 ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp  
CCAGAAATACGACTTGGAGGCTATAACATCATGCTCCTCAAACGTGTAGTCAGCGCCACGAC  
GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTGACAGTCAGCGGGTGCTG

4637 SACI,

4682 GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla  
GGCGCTGGAAAGAGGGTCTACTACCTCACCGTGACCCCTACAACCCCCCTCGCGAGAGCT  
CCCGCACCTTCTCCCAGATGATGGAGTGGGACTGGGATGTTGGGGAGCGCTCTCGA

4731 NRUI,

4742 AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe  
GCGTGGGAGACAGCAAGACACACTCCAGTCATTGCTGGCTAGGCAACATAATCATGTTT  
CGCACCCCTGTGTTCTGTGAGGTCAAGTAAAGGACCGATCCGTTGATTAGTACAAA

4802 AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla  
GCCCCACACTGTGGCGAGGATGATACTGATGACCCATTCTTAGCGTCCTTATAGCC  
CGGGGGTGTGACACCCGCTCCTACTATGACTACTGGTAAAGAAATCGCAGGAATATCGG

4806 PFLM1, 4807 DRA3,

ArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu

# FIGURE 21 - Page

4862 AGGGACCAGCTGAACAGGCCCTCGATTGCGAGATCTACGGGGCTGCTACTCCATAGAA  
 TCCCTGGTCGAACCTGTCCGGGAGCTAACGCTCTAGATGCCCGGACGATGAGGTATCTT  
 ^  
 4893 BGL2,  
 ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis  
 4922 CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTCACTCCAC  
 GGTGACCTAGATGGAGGTTAGTAAGTTCTGAGGTACCGGAGTCGCGTAAAGTGAGGTG  
 ^  
 4954 NCOI,  
 SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro  
 4982 AGTTACTCTCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAAACCTGGGGTACCG  
 TCAATGAGAGGTCCACTTAGTTATCCCACCGGGTACGGAGTCTTGAAACCCATGGC  
 ^  
 5015 SPHI, 5035 KPNI,  
 ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly  
 5042 CCCTGCGAGCTTGGAGACACCAGGGCCGGAGCGTCCCGCCTAGGCTCTGGCCAGAGGA  
 GGGAACGCTCGAACCTCTGTGGCCGGGCCTCGCAGGCGCGATCCGAAGACGGTCTCCT  
 ^  
 5064 APAI, 5091 BALI,  
 GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys  
 5102 GGCAGGGCTGCATATGTGGCAAGTACCTCTTCAACTGGCAGTAAGAACAAAGCTAAA  
 CCGTCCCGACGGTATACACCGTTATGGAGAAGTTGACCCGTATTCTGTTGAGTT  
 ^  
 5113 NDEI,  
 LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr  
 5162 CTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTCACGGCTGGCTAC  
 GAGTGAGGTTATGCCGGCGACCGGTGACCTGAACAGGCCGACCAAGTGCCGACCGATG  
 ^ ^ ^  
 5174 NOTI, 5175 EAG1 XMA3, 5182 BALI, 5186 PVU2,  
 SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys  
 5222 AGCGGGGGAGACATTATCACAGCGTGTCTCATGCCGGCCCGCTGGATCTGGTTTGC  
 TCGCCCCCTGTAAATAGTGTGCACAGAGTACGGGCCGGCGACCTAGACCAAAACG  
 ^  
 5240 DRA3,  
 LeuLeuLeuAlaAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgMetSerThrAsn  
 5282 CTACTCCTGCTGCTGCAGGGTAGGCATCTACCTCTCCCAACCGAATGAGCACGAAT  
 GATGAGGACGACGACGTCCCCATCCGTAGATGGAGGAGGGTTGGCTTACTCGTGCTTA  
 ^  
 5295 PSTI,  
 ProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGlnAspValLysPhe  
 5342 CCTAAACCTCAAAGAAAGACCAAACGTAACACCAACCGGGCCGCAGGACGTCAAGTTC  
 GGATTTGGAGTTCTTCTGGTTGCATTGTGGTTGGCCGGCGTCTGCAGTTCAAG  
 ^ ^ ^ ^  
 5380 NOTI, 5381 EAG1 XMA3, 5390 AAT2, 5401 SMA1 XMA1,  
 ProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeu  
 5402 CGGGGTGGCGGTCAAGATCGTTGGAGTTACTTGTGGTTGGCCGGCGTCTGCAGTTAGATTG  
 GGCCCACCGCCAGTCTAGCAACCACCTCAAATGAACAAACGGCGCGTCCCCGGATCTAAC  
 ^

# FIGURE 21 - Page 10

5449 APAI,

5462 GlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGlyArgArgGlnPro  
GGTGCGCGCAGAGAAAGACTTCCGAGCGGTGCAACCTCGAGGTAGACGTCAGCCT  
CCACACGCGCGCTGCTCTTCTGAAGGCTGCCAGCGTTGGAGCTCCATCTGCAGTCGGA

5467 BSSH2, 5478 XMNI, 5502 XHOI, 5511 AAT2,

5522 IleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpPro  
ATCCCCAAGGCTCGTCGGCCCGAGGGCAGGACCTGGGCTCAGCCCGGGTACCCCTGGCCC  
TAGGGGTTCCGAGCAGCCGGCTCCGCTGGACCCGAGTCGGGCCATGGGAACCGGG

5548 ALWN1, 5558 ESP1, 5564 SMAI XMAI, 5568 KPNI,

5582 LeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArg  
CTCTATGGCAATGAGGGCTCGGGGTGGCGGGATGGCTCCTGTCTCCCCGTGGCTCTCGG  
GAGATACCGTTACTCCCGACGCCACCCGCCCTACCGAGGACAGAGGGCACCGAGAGCC

5642 ProSerTrpGlyProThrAspProArgArgSerArgAsnLeuGlyLysValIleAsp  
CCTAGCTGGGGCCCCACAGACCCCCGGCGTAGGTCGCGCAATTGGTAAGGTATCGAT  
GGATCGACCCGGGTGTCGGGGCCGATCCAGCGCTAAACCCATTCCAGTAGCTA

5650 APAI, 5696 CLAI,

5702 ThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProLeuValOC AM  
ACCCTTACGTGCGGCTTCGCCGACCTCATGGGTACATACCGCTCGTCTAATAGTCGAC  
TGGGAATGCACGCCGAAGCGGCTGGAGTACCCATGTATGGCGAGCAGATTATCAGCTG

5724 HGIE2, 5755 SALI,

## FIGURE 22 - Page 1

Met Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn  
2 AGCTTACAAAACAAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC  
TCGAATGTTTGTGTTACCGACGTATACGTCGAGTCCCAGATATTCCACGATCATGAGTTG  
^  
1 HIND3, 24 NDEI, 52 SCAI,

Pro Ser Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr Met Ser Lys Ala His Gly Ile Asp  
62 CCCTCTGTTGCTGCAACACTGGGCTTGGTACATGTCCAAGGCTCATGGGATCGAT  
GGGAGACAAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCCTAGCTA  
^  
116 CLAI,

Pro Asn Ile Arg Thr Gly Val Arg Thr Ile Thr Thr Gly Ser Pro Ile Thr Tyr Ser Thr  
122 CCTAACATCAGGACCGGGGTGAGAACAAATTACCACTGGCAGCCCCATCACGTACTCCACC  
GGATTGTAGTCCTGGCCCCACTCTGTTAATGGTGACCGTCGGGGTAGTGCATGAGGTGG

Tyr Gly Lys Phe Leu Ala Asp Gly Gly Cys Ser Gly Gly Ala Tyr Asp Ile Ile Ile Cys  
182 TACGGCAAGTTCTTGCCGACGGCGGGTGCTCGGGGGCGCTTATGACATAATAATTGT  
ATGCCGTTCAAGGAACGGCTGCCGCCACGAGCCCCCGCGAATACTGTATTATAAACAA

Asp Glu Cys His Ser Thr Asp Ala Thr Ser Ile Leu Gly Ile Gly Thr Val Leu Asp Gln  
242 GACGAGTGCCACTCCACGGATGCCACATCCATCTGGCATTGGCAGTGTCTGACCAA  
CTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCCGTAACCGTGACAGGAACGGTT

Ala Glu Thr Ala Gly Ala Arg Leu Val Val Leu Ala Thr Ala Thr Pro Pro Gly Ser Val  
302 GCAGAGACTGCAGGGGGCGAGACTGGTGTGCTGCCACCGCCACCCCTCCGGCTCCGTC  
CGTCTCTGACGCCCGCTCTGACCAAACACGAGCGGTGGCGGTGGGGAGGCCAGGGCAG  
^  
303 ALWN1,

Thr Val Pro His Pro Asn Ile Glu Glu Val Ala Leu Ser Thr Thr Gly Glu Ile Pro Phe  
362 ACTGTGCCCATCCAAACATCGAGGAGGTTGCTCTGTCCACCACCGGAGAGATCCCTTT  
TGACACGGGGTAGGGTTGTAGCTCCTAACGAGACAGGTGGTGGCCTCTAGGGAAAA

Tyr Gly Lys Ala Ile Pro Leu Glu Val Ile Lys Gly Gly Arg His Leu Ile Phe Cys His  
422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGAGACATCTCATCTGTGTCAT  
ATGCCGTTCCGATAGGGGGAGCTCATTAGTCCCCCCTCTGTAGAGTAGAAGACAGTA

## FIGURE 22 - Page 2

482 SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal  
 TCAAAGAAGAGTGCAGCAACTGCCGCAGCTGGTCGCATTGGCATCAATGCCGTG  
 AGTTCTTCTTCACGCTGCTTGAGCGCGTTCGACCAGCGTAACCGTAGTTACGGCAC

542 AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal  
 GCCTACTACCGCGGTCTTGACGTGTCGTACCCGACCAGCGCGATGTTGTCGTG  
 CGGATGATGGGCCAGAACAGCACAGGCAGTAGGGCTGGTCGCCCTACAACAGCAC  
 550 SAC2, 560 DRD1,

602 AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn  
 GCAACCGATGCCCTCATGACCGGCTATACCGCGACTTCGACTCGGTGATAGACTGCAAT  
 CGTGGCTACGGGAGTACTGGCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA

615 BSPH1,

662 ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle  
 ACGTGTGTCACCCAGACAGTCGATTTCAGCCTGACCTACCTCACCATTGAGACAATC  
 TGCACACAGTGGTCTGTCAGCTAAAGTCGGAAGTGGAGTGGTAACCTGTTAG

722 ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys  
 ACGCTCCCCAAGATGCTCTCCCGACTCAACGTCGGGCAGGACTGGCAGGGGGAAAG  
 TGCGAGGGGGTCTACGACAGAGGGCGTGAGTGCAGCCCCCTGACCGTCCCCCTTC

782 ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer  
 CCAGGCATCTACAGATTGTGGCACCGGGGAGCGCCCTCCGGCATGTTGACTCGTCC  
 GGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGG

816 BGLI, 833 DRD1,

842 ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr  
 GT CCTCTGTGAGTGTATGACG CAGGCTGTGCTTGGTATGAGCTACGCCGCCGAGACT  
 CAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCAGGGCGGCTCTGA

881 SAC1,

902 ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu  
 ACAGTTAGGCTACGAGCGTACATGAAACACCCCGGGCTTCCGTGTGCCAGGACCATCTT  
 TGTCAATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGCCTGGTAGAA

931 SMA1 XMA1,

962 GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln  
 GAATTTGGGAGGGCGTCTTACAGGCCTCACTCATATAGATGCCACTTCTATCCCAG  
 CTTAAACCCCTCCCGAGAAATGTCCGGAGTGAAGTATCTACGGGTGAAAGATAGGGTC

985 STUI,

1022 ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla  
 ACAAAAGCAGAGTGGGAGAACCTCCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT  
 TGTTCTCGTCAACCCCTTGGAGGAATGGACCATCGCATGGTTCGGTGGCACACGCGA

1069 DRA3,

1082 ArgAlaGlnAlaProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys  
 AGGGCTCAAGCCCCCTCCCCCATCGTGGGACCAGATGTGGAAGTGTGATTGATTGCGCTCAAG

## FIGURE 22 - Page 5

TCCCGAGTTGGGGAGGGGGTAGCACCTGGTCTACACCTCACAAACTAAGCGGAGTTC

1142 ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle  
 CCCACCCCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTCAGAATGAAATC  
 GGGTGGGAGGTACCCGGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTAG  
 ^  
 1150 NCOI,

1202 ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal  
 ACCCTGACGCACCCAGTCACCAAATACATCATGACATGCATGTCGGCCGACCTGGAGGTC  
 TGGGACTGCGTGGGTAGTGGTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG  
 ^ ^ ^ ^  
 1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,

1262 ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu  
 GTCACGAGCACCTGGGTGCTCGTGGCGCGTCTGGCTGCTTGGCCGCGTATTGCCTG  
 CAGTGCTCGTGGACCCACGAGCAACGCCGCAGGACCGACGAAACCGGCGATAACGGAC

1322 SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle  
 TCAACAGGCTGCGTGGTCATAGTGGCAGGGTGTCTGTCCGGGAAGCCGGCAATCATA  
 AGTTGTCCGACGCACCAGTATCACCCGTCCCGAGCAGAACAGGCCCTCGGCCGTTAGTAT  
 ^  
 1369 NAEI,

1382 ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu  
 CCTGACAGGGAAAGTCCTCTACCGAGAGTTCGATGAGATGGAAGAGTGTCTCAGCACTTA  
 GGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCACCTCTCACGAGAGTCGTGAAT  
 ^  
 1385 DRD1,

1442 ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu  
 CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTCGGCCTC  
 GGCATGTAGCTCGTCCCTACTACGAGCGGCTCGTCAAGTTCGTCTCCGGGAGCCGGAG

1502 LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln  
 CTGCAGACCGCGTCCCGTCAGGCAGAGGTTATGCCCTGCTGTCCAGACCAACTGGCAA  
 GACGTCTGGCGCAGGGCAGTCGTCAGGTTGACCGTT  
 ^ ^  
 1502 PSTI, 1507 TTH3I,

1562 LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu  
 AAACTCGAGACCTTCTGGCGAAGCATATGTGAACTTCATCAGTGGATACAATACTTG  
 TTTGAGCTCTGAAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCATGTTATGAAC  
 ^ ^  
 1565 XHOI, 1586 NDEI,

1622 AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla  
 GCAGGGCTTGTCAACGCTGCTGGTAACCCCGCCATTGCTTATTGATGGCTTACAGCT  
 CGCCCGAACAGTTGCGACGGACCATTGGGGCGGTAAACGAAGTAACACGAAAATGTCGA  
 ^ ^  
 1643 BSTE2, 1677 ALWN1 PVU2,

1682 AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp  
 GCTGTCACCAGCCCACCAACCCTAGCCAAACCCCTCCTCTCAACATATTGGGGGGTGG  
 CGACAGTGGTCGGTGATTGGTGTACGGTTGGGAGGAGAGTTGTATAACCCCCCACC

**FIGURE 22 - Page 4**

1742 ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla  
 GTGGCTGCCAGCTGCCGCCCCGGTGCCGCTACTGCCTTGTGGCGCTGGCTTAGCT  
 CACCGACGGGTCGAGCGGGGGCCACGGCGATGACGGAAACACCCGCGACCGAATCGA

1794 ESP1,

1802 GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr  
 GGCGCCGCATCGCAGTGTGGACTGGGAAGGTCCTCATAGACATCCTGCAGGGTAT  
 CCGCGCGGTAGCCGTACAACCTGACCCCTTCAGGAGTATCTGTAGGAACGTCCCATA

1802 KAS1 NARI,

1862 GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer  
 GGCGGGGCGTGGCGGGAGCTCTGTGGCATTCAAGATCATGAGCGGTGAGGTCCCCTCC  
 CCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTCTAGTACTCGCCACTCCAGGGGAGG

1878 SAC1, 1899 BSPH1,

1922 ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly  
 ACGGAGGACCTGGTCAATCTACTGCCGCCATCCTCTGCCCGAGCCCTCGTAGTCGGC  
 TGCCTCCTGGACCAGTTAGATGACGGCGTAGGAGAGCGGGCTCGGGAGCATCAGCCG

1928 TTH3I,

1982 ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp  
 GTGGTCTGTGCAGCAATACTGCGCCGCACGTTGGCCCGGGCAGGGGGCAGTGCAGTGG  
 CACCAGACACGTCGTTATGACGCGGCCGTGCAACCAGGGCCGCTCCCCGTACGTCACC

2004 NAEI, 2017 SMAI XMAI,

2042 MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal  
 ATGAACCGGCTGATAGCCTTCGCCCTCCGGGGAACATGTTCCCCCACGCACACTACGTG  
 TACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTGGTACAAAGGGGTGCGTGATGCAC

2067 SMAI XMAI, 2093 DRA3,

2102 ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln  
 CCGGAGAGCGATGCAGCTGCCGCCTACTGCCATACTCAGCAGCCTCACTGTAACCCAG  
 GGCCTCTCGCTACGTCGACGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGGTC

2115 PVU2, 2159 ALWN1,

2162 LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer  
 CTCCTGAGGCGACTGCACAGTGGATAAGCTGGAGTGTACCACTCCATGCTCCGGTTCC  
 GAGGACTCCGCTGACGTGGTACCTATTGAGCCTCACATGGTGAGGTACGAGGCCAAGG

2164 MST2, 2220 ECON1,

2222 TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu  
 TGGCTAAGGGACATCTGGACTGGATATGCGAGGTGTTGAGCGACTTTAACGACCTGGCTA  
 ACCGATTCCCTGTAGACCTTGACCTATAACGCTCCACAACTCGCTGAAATTCTGGACCGAT

2282 LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr  
 AAAGCTAACGCTCATGCCACAGCTGCCTGGATCCCCTTGTGTCCTGCCAGCGCGGGTAT  
 TTTCGATTGAGTACGGTGTGACGGACCCCTAGGGGAAACACAGGACGGTCGCGCCCATA

2285 ESP1, 2300 PVU2, 2310 BAMHI,

## FIGURE 22 - Page 5

2342 LysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAlaGluIle  
 AAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC  
 TTCCCCCAGACCGCTCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAG

2402 ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet  
 ACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCTAGGACCTGCAGGAACATG  
 TGACCTGTACAGTTTGCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCCCTGTAC

^ ^ ^

2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,

2462 TrpSerGlyThrPheProIleAsnAlaTyrThrGlyProCysThrProLeuProAla  
 TGGAGTGGGACCTCCCCATTAAATGCCCTACACCACGGGCCCCGTACCCCCCTTCCCTGCG  
 ACCTCACCCCTGGAAGGGGTAAATTACGGATGTGGTCCCCGGGACATGGGGGAAGGACGC

^

2480 ASE1, 2497 APA1,

2522 ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln  
 CCGAACTACACGTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATAAGGCAG  
 GGCTTGATGTGCAAGCGCGATACCTCCCACAGACGTCTCCTATGCACCTCTATTCCGTC

^

2553 PSTI,

2582 ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln  
 GTGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTAAATGCCGTGCCAG  
 CACCCCCCTGAAGGTGATGCACTGCCCATACTGATGACTGTTAGAATTACGGGCACGGTC

^

2594 DRA3,

2642 ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro  
 GTCCCCATCGCCCGAATTTCACAGAATTGGACGGGTGCGCCTACATAGGTTGCGCCC  
 CAGGGTAGCGGGCTTAAAAAGTGTCTAACCTGCCAACCGGGATGTATCCAAACGCGGG

2702 ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro  
 CCCTGCAAGCCCTTGCTCGGGAGGGAGGTATCAGAGTAGGACTCCACGAATACCCG  
 GGGACGTTCGGAACGACGCCCTCCTCCATAGTAAGTCTCATCCTGAGGTGCTTATGGC

^

2757 HGIE2,

2762 ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu  
 GTAGGGTCGCAATTACCTTGCAGGCCGAACCGACGTGGCGTGTGACGTCCATGCTC  
 CATCCCAGCGTTAATGGAACGCTCGGGCTTGGCCTGCACCGGACAACTGCAGGTACGAG

^

2809 AAT2,

2822 ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro  
 ACTGATCCCTCCCATAAACAGCAGAGGCAGGCCGGCGAAGGTTGGCGAGGGGATCACCC  
 TGACTAGGGAGGGTATATTGTCGTCTCCGCCGGCCCGCTTCCAACCGCTCCCTAGTGGG

^

2850 EAG1 XMA3,

2882 ProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAlaThrCys  
 CCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCAAGGCAACTTGC  
 GGGAGACACCGGTCGAGGAGCCGATCGGTGCGATAGGCGAGGTAGAGAGAGTTCCGTTGAACG

^ ^

2889 BALI, 2903 NHEI,

**FIGURE 22 - Page 6**

2942 ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln  
 ACCGCTAACCATGACTCCCCTGATGCTGAGCTCATAGAGGCCAACCTCTATGGAGGCAG  
 TGGCGATTGGTACTGAGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACCTCCGTC  
 ^ ^

2966 ESP1, 2969 SACI,

3002 GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer  
 GAGATGGCGGCAACATCACCAAGGGTTGAGTCAGAAAACAAAGTGGTATTCTGGACTCC  
 CTCTACCCGCCGTTGAGTGGTCCCAACTCAGTCTTTGTTCACCACTAACGACCTGAGG

3062 PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu  
 TTCGATCCGCTTGTGGCGGAGGAGCAGCGGGAGATCTCCGTACCCGCAGAAATCTG  
 AAGCTAGGCGAACACCGCCTCCTGCTCGCCCTAGAGGCATGGCGTCTTAGGAC  
 ^

3096 BGL2,

3122 ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro  
 CGGAAGTCTCGGAGATTGCCCAGGCCCTGCCGTTGGCGCGGCCGGACTATAACCCC  
 GCCTTCAGAGCCTAAAGCGGTCCGGACGGGCAAACCCGCGCCGGCTGATATTGGGG  
 ^ ^

3143 ALWN1, 3164 EAG1 XMA3,

3182 ProLeuValGluThrTrpLysProAspTyrGluProProValValHisGlyCysPro  
 CCGCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACCACCTGTGGTCCATGGCTGCCCG  
 GGCGATCACCTCTGCACCTTTGGCTGATGCTTGGACACCAGGTACCGACGGC  
 ^ ^

3217 HGIE2, 3229 NCOI,

3242 LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu  
 CTTCCACCTCAAAGTCCCCTCCTGTGCCCTCGAACGGAGGAGCAGCGGGAGCCTTCG  
 GAAGGTGGAGGTTCAAGGGAGGACACGGAGGCGGAGCCTTCGCTGCCACCAGGAG

3302 ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer  
 ACTGAATCAACCCTATCTACTGCCCTGGCGAGCTGCCACCAGAACGCTTGGCAGCTCC  
 TGACTTAGTTGGATAGATGACGGAACCGGCTGAGCGGTGGTCTCGAAACCGTCGAGG  
 ^ ^

3332 SACI, 3346 HIND3,

3362 SerThrSerGlyIleThrGlyAspAsnThrThrSerSerGluProAlaProSerGly  
 TCAACTTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCGCCCTCTGGC  
 AGTTGAAGGCCGTAATGCCCGTGTATGCTGTTAGGAGACTCGGGCGGGAAAGACCG

3422 CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro  
 TGCCCCCCCCGACTCCGACGCTGAGTCCTATTCCCTCATGCCCGCCCTGGAGGGGGAGCCT  
 ACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGACCTCCCCCTCGGA  
 ^

3437 EAM11051,

3482 GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu  
 GGGGATCCGGATCTAGCGACGGGTATGGTCACGGTCAGTAGTGAGGCCAACCGCGAG  
 CCCCTAGGCCTAGAATCGCTGCCAGTACCAAGTGGCCAGTCATCACTCCGGTTGCGCCTC  
 ^ ^

3484 BAMHI, 3485 BSAB1, 3487 BSPE1,

3542 AspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla  
 GATGTCGTGTGCTGCTCAATGTCCTACTCTGGACAGGCGCACTCGTCACCCCGTGC  
 CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCCGCTGAGCAGTGGGCACCGC

## FIGURE 22 - Page 7

3589 DRA3, 3600 SAC2,

3602 AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn  
GCGGAAGAACAGAAACTGCCATCAATGCACTAAGCAACTCGTGCTACGTACCGACAAAT  
CGCCTTCTGCTTTGACGGGTAGTTACGTGATTGAGCAACGATGCAGTGGTGTAA

3611 ALWN1, 3655 PFLM1,

3662 LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp  
TTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAGAAAGTCACATTGAC  
AACACACATAAGGTGGTGGAGTGCGTCACGAACGGTTCCGTCTTCAGTGTAAACTG

3681 DRA3,

3722 ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla  
AGACTGCAAGTCTGGACAGCCATTACCAAGGACGTACTCAAGGAGGTAAAGCAGCGCG  
TCTGACGTTCAAGACCTGTCGTAATGGCCTGCATGAGTCCCTCAATTTCGTCGCCGC

3782 SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis  
TCAAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGAGCCTGACGCCACAC  
AGTTTCACTTCCGATTGAACGATAGGCATCTCCTCGAACGTCGGACTGCGGGGGTGTG

3816 HIND3,

3842 SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla  
TCAGCCAAATCCAAGTTGGTATGGGCAAAAGACGTCCGTGAGCCTGACGCCAGAAAGGCC  
AGTCGGTTAGGTTCAAACCAATACCCGTTCTGAGGCAACGGTACGGTCTTCGG

3875 AAT2, 3890 BGLI,

3902 ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp  
GTAACCCACATCAAACCTCCGTGTGGAAAGACCTCTGGAAAGACAATGTAACACCAATAGAC  
CATTGGGTGTAGTTGAGGCACACCTTCTGGAAGACCTCTGTTACATTGGTTATCTG

3962 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyArgLys  
ACTACCACATGGCTAAAGAACGAGGTTCTGCGTTAGCCTGAGAAGGGGGTGTAAAG  
TGATGGTAGTACCGATTCTGCTCAAAGACGCAAGTCGGACTCTCCCCCAGCATTG

4022 ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu  
CCAGCTCGTCTACGTGTCCCCGATCTGGCGTGCCTGCGAAAGATGGCTTG  
GGTCGAGCAGAGTAGCACAGGGCTAGACCGCACGCGCACAGCTTCTACCGAAAC

4082 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr  
TACGACGTGGTTACAAAGCTCCCTGGCGTGATGGAAAGCTCCTACGGATTCAAATAC  
ATGCTGCACCAATGTTGAGGGAAACCGGCACCTCGAGGATGCCAAGGTTATG

4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet  
TCACCAGGACAGCGGGTTGAATTCTCGTGCAAGCGTGGAAAGTCCAAGAAAACCCAATG  
AGTGGCTCTGTCGCCAACTTAAGGAGCAGTTCGACCTTCAGGTTCTTGGGTTAC

4160 ECORI,

4202 GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr  
GGGTTCTCGTATGATAACCGCTGCTTGACTCCACAGTCACTGAGAGCGACATCCGTACG  
CCCAAGAGCATACTATGGCGACGAAACTGAGGTGTCAGTGAATCTCGCTGTAGGCATGC

## FIGURE 22 - Page 8

4229 DRD1, 4236 ALWN1,

4262 GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer  
GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCCAAGCCCCGTGGCCATCAAGTCC  
CTCCTCGTTAGATGGTACAACACTGGAGCTGGGGTCGGGCGCACCGTAGTTAGG

4301 BGLI, 4308 BALI,

4322 LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly  
CTCACCGAGAGGCTTATGTTGGGGCCCTTACCAATTCAAGGGGGAGAACTGCGGC  
GAGTGGCTCTCCGAAATACAACCCCCGGGAGAATGGTTAAGTCCCCCTTTGACGCCG

4345 APA1,

4382 TyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCys  
TATCGCAGGTGCCGCGCAGCGCGTACTGACAAGTAGCTGTGGTAACACCCTCACTTGC  
ATAGCGTCCACGGCGCGCTCGCCGATGACTGTTGATCGACACCATTGTGGAGTGAACG

4442 TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal  
TACATCAAGGCCGGGCAGCCTGTCGAGCCGCAAGGGCTCCAGGACTGCACCATGCTCGTG  
ATGTAGTTCCGGGCCGTCGGACAGCTGGCGTCCGAGGTCTGACGTGGTACGAGCAC

4452 SMA1 XMA1,

4502 CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer  
TGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGTCCAGGAGGACGCGCGAGC  
ACACCGCTGCTGAATCAGCAATAGACACTTCCGCCCCCAGGTCTCCTGCGCCGCTCG

4508 DRD1, 4511 TTH3I,

4562 LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln  
CTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCTGGGACCCCCCACAA  
GACTCTCGGAAGTGCCTCCGATACTGGTCCATGAGGCGGGGGGACCCCTGGGGGGTGT

4622 ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp  
CCAGAATAACGACTGGAGGCTATAACATCATGCTCCTCCAACGTGTAGTCGCCACGAC  
GGTCTTATGCTGAACCTCGAGTATTGTAGTACCGAGGAGGTGCACAGTCAGCGGGTGCTG

4637 SAC1,

4682 GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla  
GGCGCTGGAAAGAGGGCTACTACCTCACCGTGACCCCTACAACCCCCCTCGCGAGAGCT  
CCCGGACCTTCTCCAGATGATGGAGTGGGACTGGGATGTTGGGGAGCGCTCTCGA

4731 NRUI,

4742 AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe  
GCGTGGGAGACAGCAAGACACACTCCAGTCATTGCTGGCTAGGCAACATAATCATGTT  
CGCACCCCTCTGCGTTCTGTGAGGTCAAGGACCGATCCGTTGATTAGTACAAA

4802 AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla  
GCCCCACACTGTGGCGAGGATGATACTGATGACCCATTCTTAGCGTCCTTATAGCC  
CGGGGGTGTGACACCCGCTCCTACTATGACTACTGGTAAAGAAATCGCAGGAATATCGG

4806 PFLM1, 4807 DRA3,

ArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu

**FIGURE 22 - Page 9**

4862 AGGGACCAGCTGAAACAGGCCCTCGATTGCGAGATCTACGGGGCTGCTACTCCATAGAA  
TCCCTGGTCGAACCTGTCCGGAGCTAACGCTCTAGATGCCCGGACGATGAGGTATCTT  
4893 BGL2,  
ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis  
4922 CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTCACTCCAC  
GGTGACCTAGATGGAGGTTAGTAAGTTCTGAGGTACCGGAGTCGCGTAAAGTGAGGTG  
4954 NCOI,  
SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro  
4982 AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAAACCTGGGGTACCG  
TCAATGAGAGGTCCACTTAGTTATCCCACCGGCGTACGGAGTCTTGAAACCCATGGC  
5015 SPHI, 5035 KPNI,  
ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly  
5042 CCCTTGCAGCTTGAGACACCAGGGCCGGAGCGTCCGCGTAGGCTCTGGCCAGAGGA  
GGGAACGCTCGAACCTCTGTGGCCGGCCTCGCAGGCGCGATCCGAAGACCGGTCTCCT  
5064 APAI, 5091 BALI,  
GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys  
5102 GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGAGTAAGAACAAAGCTAAA  
CCGTCCCGACGGTATACACCGTTATGGAGAAGTTGACCCGTATTCTGTTGAGTT  
5113 NDEI,  
LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr  
5162 CTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGCGTGGTACGGCTGGCTAC  
GAGTGAGGTTATGCCGGCGACGGTCGACCTAACAGGCCGACCAAGTGCCGACCGATG  
5174 NOTI, 5175 EAG1 XMA3, 5182 BALI, 5186 PVU2,  
SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys  
5222 AGCAGGGGAGACATTTATCACAGCGTGTCTCATGCCGGCCCGCTGGATCTGGTTTGC  
TCGCCCCCTCTGTAATAGTGTGCACAGAGTACGGGCCGGGCGACCTAGACCAAAACG  
5240 DRA3,  
LeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgMetSerThrAsn  
5282 CTACTCCTGCTTGCAGGGTAGGCATCTACCTCCTCCCCAACCGAATGAGCACGAAT  
GATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGTTGGTTACTCGTGCTTA  
5295 PSTI,  
ProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGlnAspValLysPhe  
5342 CCTAAACCTCAAAGAAAGACCAACCGTAAACACCAACCGGCCGCAGGACGTCAAGTTC  
GGATTGGAGTTCTTCTGGTTGCATTGTGGTTGGCCGCCGGCGTCTGCAGTTCAAG  
5380 NOTI, 5381 EAG1 XMA3, 5390 AAT2, 5401 SMA1 XMA1,  
ProGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeu  
5402 CCGGGTGGCGGTACAGATCGTGGTGGAGTTACTTGTGGCCGCGCAGGGGCCCTAGATTG  
GGCCCACCGCCAGTCTAGCAACCACCTCAAATGAACAAACGGCGCGTCCCCGGGATCTAAC

FIGURE 22 - Page 10

5449 APAI,

5462 GlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGlyArgArgGlnPro  
GGTGTGCGCGCGACGAGAAAGACTTCCGAGCGGTGCAACCTCGAGGTAGACGTCAGCCT  
CCACACGCGCGCTGCTCTTCTGAAGGCTGCCAGCGTTGGAGCTCCATCTGCAGTCGGA

5467 BSSH2, 5478 XMNI, 5502 XHOI, 5511 AAT2,

5522 IleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpPro  
ATCCCCAAGGCTCGTCGGCCCGAGGGCAGGACCTGGCTCAGCCCGGGTACCCCTGGCCC  
TAGGGGTTCCGAGCAGCCGGCTCCCGCTGGACCCGAGTCGGGCCATGGGAACCGGG

5548 ALWN1, 5558 ESP1, 5564 SMAI XMAI, 5568 KPNI,

5582 LeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArg  
CTCTATGGCAATGAGGGCTCGGGGTGGCGGGATGGCTCCTGTCTCCCCGTGGCTCTCGG  
GAGATAACGTTACTCCCGACGCCAACCGCCCTACCGAGGACAGAGGGCACCGAGAGCC

5642 ProSerTrpGlyProThrAspProArgArgSerArgAsnLeuGlyLysValIleAsp  
CCTAGCTGGGGCCCCACAGACCCCGGCGTAGGTCGCGCAATTGGGTAAAGGTATCGAT  
GGATCGACCCCGGGGTGTCTGGGGCCGCATCCAGCGCGTTAACCCATTCCAGTAGCTA

5650 APAI, 5696 CLAI,

5702 ThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProLeuValGlyAlaProLeu  
ACCCTTACGTGCGGCTTCGCCGACCTCATGGGTACATACCGCTCGTGGCGCCCCCTTT  
TGGGAATGCACGCCGAAGCGGCTGGAGTACCCATGTATGGCGAGCAGCCGCGGGGAGAA

5724 HGIE2, 5750 KAS1 NARI, 5756 ECON1,

5762 GlyGlyAlaAlaArgAlaOC AM  
GGAGGCGCTGCCAGGGCCTAATAGTCGAC  
CCTCCGCGACGGTCCCGGATTATCAGCTG

5785 SALI,

**FIGURE 23**

